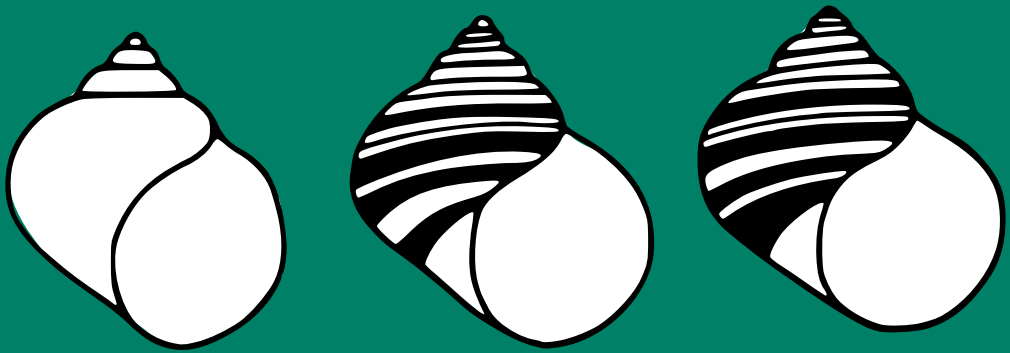
A. Sosinov, Y. Altoukhov

GÉNÉTIQUE, ÉVOLUTION ET ENVIRON-NEMENT



Éditions Mir Moscou

СБОРНИК СТАТЕЙ ПО СРАВНИТЕЛЬНОЙ ГЕНЕТИКЕ ПОПУЛЯЦИЙ

Под редакцией А. А. СОЗИНОВА
академика АН УССР

Составитель Ю. П. АЛТУХОВ
д-р биологических наук

GÉNÉTIQUE, ÉVOLUTION ET ENVIRONNEMENT

Sous la direction de A. SOSINOV
Edition établie par Y. ALTOUKHOV

ÉDITIONS MIR · MOSCOU

Traduit du russe
par E. BRONINA

À NOS LECTEURS

Les Editions Mir vous seraient très reconnaissantes de bien vouloir leur communiquer votre opinion sur le contenu de ce livre, sa traduction et sa présentation, ainsi que toute autre suggestion.

Notre adresse : Editions Mir, 2, Pervi
Rijski péréoulok, Moscou,
I-110, GSP, U.R.S.S.

На французском языке

© Traduction française • Editions Mir • 1985

AVANT-PROPOS

Parmi les courants de la science, celui des recherches réunies dans le cadre du problème « L'homme et le Milieu » s'avance aujourd'hui au premier plan. Une telle tendance est dictée par la vie même, car, de nos jours, l'action de l'Homme sur la Nature a pris une ampleur qu'elle n'avait jamais eue de toute l'histoire connue.

Cependant, la Nature et la Société représentent un système dynamique formant un tout et, par conséquent, les changements survenant dans le monde de la Nature ne manquent pas de se répercuter sur la nature de l'homme même. Les phénomènes nocifs provoqués par des troubles de l'équilibre écologique peuvent retentir aussi bien sur les êtres actuellement vivants que sur les générations futures pour des causes diverses: alourdissement du fardeau génétique, appauvrissement de la diversité génétique des populations, modification de la structure démographique qui s'était constituée au cours de l'histoire, etc.

Les aspects génétiques du problème « L'Homme et la Biosphère » ont pris aujourd'hui tant d'acuité qu'on leur accorde une attention particulière dans le programme de l'O.N.U. sur l'environnement. Ce programme affirme nettement que les mesures visant à conserver et à protéger la biosphère naturelle assurent une diversité génétique constante pour les populations humaines et sont la seule garantie de l'existence de l'Homme même.

Les questions relatives au maintien de l'hérédité de l'homme dans les conditions d'un milieu qui change vite attirent vivement l'attention de l'Organisation mondiale de la Santé. Dans l'Acte final de la

Conférence sur la sécurité et la coopération en Europe, signé le 1 août 1975 à Helsinki, les Etats participants ont déclaré que la protection et l'amélioration de l'environnement, ainsi que l'utilisation rationnelle des ressources naturelles dans l'intérêt des générations actuelles et futures sont des tâches qui ont le plus d'importance pour le bien-être des peuples et le développement économique de tous les pays. L'Acte indique parmi les domaines les plus importants de la coopération internationale : la prévision des mutations génétiques possibles dans les règnes animal et végétal par suite de la pollution et l'estimation de l'influence des degrés de pollution et de dégradation du milieu sur la santé humaine.

Le but du présent recueil est de faire connaître aux biologistes français les recherches poursuivies dans le cadre du problème « L'Homme et la Biosphère » à l'Institut de Génétique générale de l'Académie des Sciences, organisation pilote dans cette branche en Union Soviétique.

Dans notre pays, l'ampleur de ce genre de travaux est appréciable et, bien entendu, ce livre n'en donne qu'une idée incomplète.

Cependant, la collection d'articles que nous avons choisis se distingue de toutes les autres recherches analogues en ce sens qu'elle représente la première approche de ce problème à partir des positions de la génétique des populations.

Le cours de l'histoire a été tel que la génétique des populations, née comme science exacte au début du siècle et vouée à l'étude des relations de conservation et de transformation de l'information héréditaire dans les générations, s'est principalement intéressée aux problèmes évolutionnistes de la biologie et a surtout centré son attention sur les principaux aspects de la dynamique génétique des populations (des générations). Un tel abord a contribué à former une théorie probabiliste de l'évolution, minutieusement mise au point, et exprimée en termes génétiques, mais a pour longtemps détourné l'attention des chercheurs du fait qu'il existe des questions clés qui seront élaborées avec succès plutôt en tenant compte des facteurs et des conditions de stabilité génétique des populations que des facteurs de leur évolution. Cet aspect du problème prédomine de nos jours où le milieu environnant change si brusquement.

Cela se comprend, car, si la pression anthropogène est ressentie par la biosphère dans son ensemble, le point d'impact des actions extérieures sont toujours et directement les populations, unités principales de notre activité pratique et forme essentielle d'existence de l'Homme même en tant qu'espèce biologique qui s'est formée au cours d'un long développement historique.

Ces dernières années, l'aspect génétique du problème « L'Homme et la Biosphère » est largement discuté dans la littérature mondiale.

Cependant, l'attention est surtout attirée par la pollution chimique de l'environnement et par la mise au point de procédés spéciaux, de systèmes de tests permettant d'identifier les mutagènes dans l'immense flot de composés chimiques divers couramment utilisés dans l'économie nationale, la médecine, la vie domestique. Sans amoindrir aucunement l'importance de ce courant des travaux, il faut pourtant indiquer que, de notre point de vue, on ne saurait, dans cette voie, trouver la clef du problème. Nous avons besoin d'une approche qui permettrait d'apprécier directement les processus génétiques se déroulant actuellement dans les populations humaines et autres.

Nous devons comprendre la spécificité de ce genre de mutations, rechercher, s'il est possible, d'en faire la prévision et élaborer des mesures pratiques visant à éliminer les effets indésirables. Parmi ceux-ci, la dynamique du fardeau génétique dans les populations humaines, l'appauvrissement du patrimoine génétique collectif des populations de plantes et d'animaux créent des problèmes qui exigent une attention spéciale.

Ces dernières années, notre collectif a effectué dans ce sens un travail dont nous présentons les résultats au jugement du lecteur français. Les articles sont groupés de façon à mettre en évidence deux lignes principales dans nos recherches dont l'une est en rapport avec la génétique humaine et l'autre, avec le patrimoine génétique collectif des populations d'espèces qui sont, actuellement, les objets d'une activité économique intense. Dans les travaux que nous proposons, qui prolongent et développent tout naturellement nos recherches antérieures, l'attention porte principalement sur la structure des populations et sur son importance pour le maintien de la

stabilité génétique de l'espèce, ainsi que sur les possibilités de résoudre cette question, offertes par de récents travaux dans le domaine de la génétique biochimique des populations. Bien que ce recueil ne donne pas une généralisation complète de toutes les données obtenues, nous espérons que la ligne générale de cette approche est assez claire et présente de l'intérêt pour nos collègues français.

Les auteurs

LA GÉNÉTIQUE DES POPULATIONS, L'ÉVOLUTION ET LE FOND GÉNIQUE DE LA BIOSPHÈRE

LA GÉNÉTIQUE BIOCHIMIQUE DES POPULATIONS ET L'ÉVOLUTION

Y. Altoukhov

Il y a plusieurs années, une hypothèse de formation des espèces réunissant l'approche typologique et l'approche de la génétique des populations du problème de l'espèce a été formulée, s'appuyant sur les recherches dans le domaine de la génétique biochimique des populations et de la génétique comparée des espèces [1, 4, 6, 14, 15].

Les dernières années ont apporté de nouvelles preuves de cette hypothèse et bien que le temps ne soit pas encore arrivé pour une synthèse définitive, il est intéressant d'aborder cette question compte tenu des élaborations théoriques et expérimentales effectuées dans les domaines correspondants de la génétique des populations et de la génétique biochimique.

Pour éviter les reproches des critiques, l'auteur tient à dire que, bien que le terme d'« évolution » figure dans le titre de cet article, il se borne à examiner seulement le problème de l'espèce et de la formation des espèces, c'est-à-dire de la microévolution, et la preuve d'une transformation évolutive, en accord complet avec la tradition, sera un écart vectorisé dans le temps (ou dans l'espace) des fréquences géniques au niveau des populations.

Il faut également souligner que, la génétique des populations ayant affaire à des formes récentes et à des intervalles de temps qui ne représentent qu'un bref instant sur l'échelle de l'évolution, les modifications génétiques accompagnant la sortie de la population hors des limites de l'espèce, ne peuvent être saisies qu'en recourant au procédé méthodique suivant: si la différence entre l'espèce et la variante n'est pas dans l'essence, mais dans le degré et que la variante soit une espèce en cours de naissance, le caractère des différences entre les espèces dans les limites du genre doit être commen-

© Издательство Наука, 1982.

surable avec celui des différences entre les populations dans les limites de l'espèce.

Compte tenu de ces remarques, nous essayerons, dans cet article, de fonder les trois thèses suivantes qui sont d'une importance fondamentale pour le déchiffrement des bases génétiques de la microévolution : 1° on ne saurait parvenir à l'idée d'évolution par l'analyse des phénomènes génétiques au niveau des populations ; 2° pour comprendre les lois de la formation des espèces, il ne suffit pas de se borner au phénomène du polymorphisme héréditaire, il faut en même temps étudier le système génétique monomorphe de l'espèce ; 3° la formation des espèces n'est pas un remaniement graduel du fond génétique de la population, c'est le résultat d'une réorganisation qualitative du système génétique monomorphe, fonctionnellement le plus important, de l'espèce.

Dans la conclusion de cet article, l'auteur examinera l'importance qu'a une approche de ce genre en raison des aspects génétiques actuellement brûlants du problème « L'Homme et la Biosphère ».

On ne saurait parvenir à l'idée d'évolution par l'analyse des phénomènes génétiques au niveau des populations

Bien que qualifiée de synthétique ou biologique, la conception moderne de l'espèce et de la formation des espèces est, par son essence, génétique, car c'est dans son cadre qu'a été réalisée la synthèse de la théorie darwinienne de la sélection naturelle et des acquis de la génétique mendélienne [24]. Pour nous en rendre compte une fois de plus, examinons le schéma proposé par Th. Dobzhansky [26] qui, sous une forme généralisée, représente les modèles aussi bien de la « vraie formation des espèces », c'est-à-dire de la scission de l'espèce originelle en deux, que de l'« évolution phylétique », c'est-à-dire de la transformation de l'espèce en tant qu'entité dans le temps (fig. 1).

Il est évident qu'en accord avec le développement des idées et des conceptions de la génétique des populations, ce ne sont pas, sur le schéma de Dobzhansky, des populations mendéliennes isolées qui divergent, mais leur ensemble, ce qu'on appelle les populations subdivisées ou, d'après notre terminologie, les systèmes de populations [14]. C'est seulement dans une telle population (S. Wright la considère souvent comme une espèce) que se réunissent les conditions favorisant le maximum de diversité génétique, ce qui lui permet de réagir le plus efficacement possible aux changements de l'environnement et de modifier en l'occurrence sa structure génétique [27, 28, 29].

Mais, cette conclusion théorique n'était suffisamment confirmée ni par des recherches expérimentales ni par des essais en champ qui restent encore souvent rattachés à des échantillons isolés ou, dans

le meilleur des cas, à certaines populations élémentaires n'ayant qu'une importance subordonnée dans la Nature. Peut-être que simplement en vigueur d'une tradition liée au grand intérêt porté au problème de l'évolution, on estimait et on estime encore aujourd'hui que si les transformations au niveau des populations élémentaires servent de témoignage direct de l'évolution qui se poursuit sous nos yeux, les possibilités de l'évolution seront d'autant plus grandes

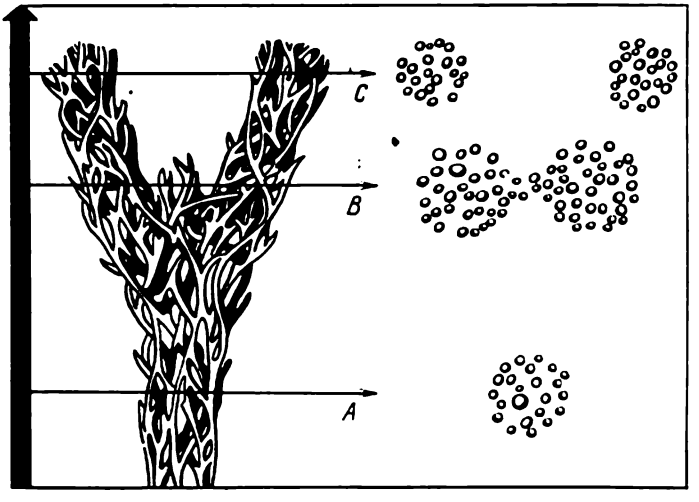


Fig. 1. Représentation schématique de la scission d'une espèce originelle (A) en deux espèces (B et C) dans le temps (flèche):
Les branches à part plus ou moins entrelacées sont des populations mendéliennes (selon Dobzhansky, 1955)

pour un ensemble de populations de ce genre et encore plus pour les espèces largement distribuées — ces véritables pionniers de l'évolution qui ne cessent d'expérimenter sur elle [44].

Cependant, si l'on ne se borne pas à l'investigation au niveau des populations élémentaires, mais si l'on étudie les populations naturelles dans leurs frontières naturelles, historiquement constituées, on peut découvrir une qualité systémique importante qui leur est propre, qu'on ne saurait faire découler avec évidence des propriétés des composantes formant leur structure: la stabilité génétique dans le temps et dans l'espace. Il s'est avéré que, dans un système de populations, le processus génétique a un déroulement stationnaire basé sur l'équilibre réciproque des pressions de facteurs aléatoires et systématiques de l'évolution sur sa structure interne. Des effets de ce genre ont été découverts, pour la première fois, en étudiant certains isolats humains [54, 55] et, ensuite, grâce aux tra-

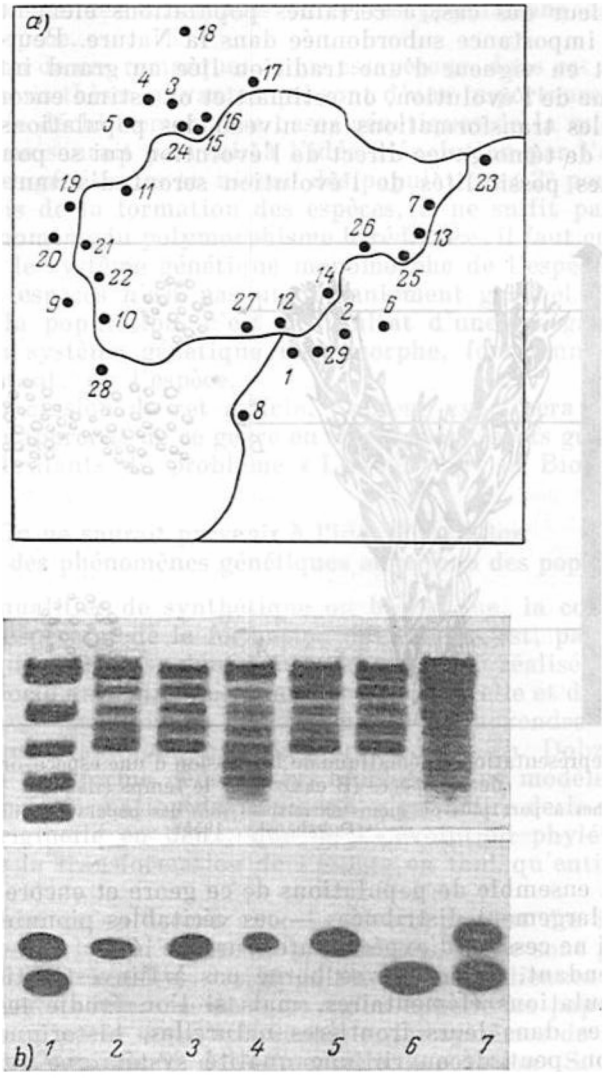


Fig. 2. Localisation spatiale des populations de frai de saumon rouge (*Oncorhynchus nerka* Walb.) (*1-29*) dans le lac Azabatchi au Kamtchatka (*a*) et polymorphisme héréditaire des enzymes lactate-déshydrogénase sérique (en haut) et phosphoglucomutase musculaire (en bas) dans les populations étudiées de saumon rouge (*b*):

Partie supérieure de la figure: *1* — homozygote B'B'; *2-3* et *5-6* — homozygotes BB; *4* et *7* — hétérozygotes BB'.
Partie inférieure de la figure: *1* et *7* — hétérozygotes AB; *2-5* homozygotes AA; *6* — homozygote BB

vaux sur la génétique biochimique et la génétique des populations des poissons [1, 4, 5, 13, 17]. Examinons en qualité d'exemple les données correspondantes pour un banc local de saumons rouges (*Oncorhynchus nerka* Walb.), salmonidés du Pacifique. Cette population isolée, qui se reproduit dans le lac Azabatchi au Kamtchatka (fig. 2), est étudiée par notre groupe depuis 8 saisons en champ en ce qui concerne deux loci polymorphes de gènes codant pour les isoenzymes de la lactate-déshydrogénase (LDH) et de la phosphoglucomutase (PGM). On peut montrer quelle importante dispersion de

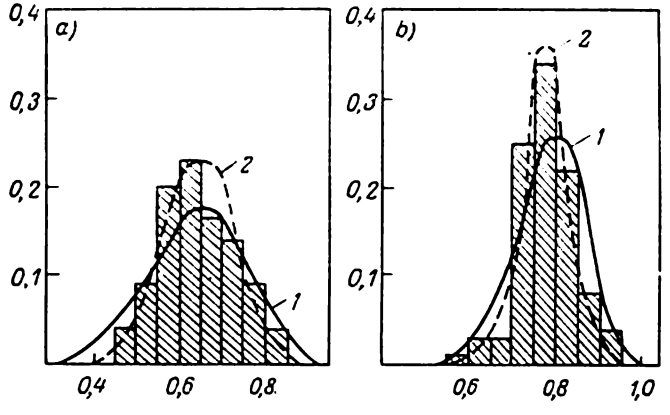


Fig. 3. Distribution des fréquences des allèles de la lactate-déshydrogénase (*a*) et de la phosphoglucomutase (*b*) dans le système de subpopulations de saumon rouge (*Oncorhynchus nerka* Walb.) du lac Azabatchi:

Les histogrammes sont les distributions empiriques; les courbes, les distributions théoriques:

$$1 - \Phi(q) = Cq^{4\bar{N}m\bar{q}-1}(1-q)^{4\bar{N}m\bar{p}-1};$$

$$2 - \Phi(q) = C(\overline{W})^{2\bar{N}}q^{4\bar{N}m\bar{p}-1}(1-q)^{4\bar{N}m\bar{p}-1}.$$

Sur l'axe des abscisses, la fréquence des allèles des loci LDH-B et PGM-A; sur l'axe des ordonnées, le nombre de subpopulations (en fractions du nombre total d'individus étudiés) avec une fréquence donnée.

Valeur des coefficients d'adaptabilité des génotypes: pour LDH $W_{BB} = 0{,}98$; $W_{BB'} = 1{,}00$; $W_{B'B'} = 0{,}97$; pour PGM $W_{AA} = 0{,}94$; $W_{AB} = 1{,}00$; $W_{BB} = 0{,}82$

fréquences des allèles est observée à l'intérieur de ce système de populations, surtout pour le locus LDH (fig. 3).

Si l'on étudie d'année en année les mêmes subpopulations de frai, on découvre une importante variabilité des fréquences des allèles non seulement dans l'espace, mais aussi dans le temps. Mais si l'on étudie le banc dans son ensemble, l'appréciant sous le rapport des valeurs moyennes des fréquences géniques et de leurs dispersions, on trouve que, malgré la variabilité dans les parties, une telle population reste génétiquement stable. Pratiquement, la fréquence moyenne du gène dans le système ne se distingue pas de la

fréquence d'équilibre trouvée par le rapport des adaptabilités des génotypes, la valeur de la dispersion intergroupe des fréquences géniques (tout au moins pour le locus PGM le plus susceptible de sélection) n'est pas élevée. Ceci engage à supposer qu'un tel système a déjà atteint son maximum d'adaptation et, par conséquent, même pour de violents changements de l'environnement, on ne saurait s'attendre à son évolution [13, 17]. En effet, pour les espèces de salmonidés ayant une structure de populations rigide, de nombreuses mesures d'acclimatation restent sans succès. C'est ainsi que nous avons montré que, malgré de nombreux transferts des œufs de saumon kéta aussi bien à l'intérieur de la région de Sakhaline et des Kouriles qu'en dehors d'elle, nous n'avons jusqu'à présent aucune preuve de l'efficacité de ces mesures [13, 14, 17, 18]. Bien plus, on a découvert que dans les conditions nouvelles la population transférée ne modifie pas sa structure génétique malgré une puissante sélection. Il est clair qu'une sélection de ce genre ne « sélectionne » pas, elle est catastrophique [18].

Il est donc évident que si nous confrontons des échantillons pris au hasard en étudiant les particularités de la différenciation génétique dans l'espace ou en faisant des coupes successives à travers la même population dans le temps, nous pouvons parvenir à conclure aux phénomènes évolutifs là où en réalité on n'a pas le droit de le faire, ou, au contraire, nous conclurons à la stabilité là où il se déroule un processus génétique vectorisé. Seule l'analyse approfondie du fond génique des populations, compte tenu de leur structure de subpopulations interne, permet une conclusion justifiée relativement au caractère du processus génétique se déroulant dans le système de populations. En appliquant cette approche à diverses espèces, on peut montrer comment, sur un fond de la variabilité de diverses subpopulations, le système dans son ensemble conserve ses traits génétiques pratiquement inchangés au fil des générations.

Auparavant, on a montré bien des exemples reflétant des processus génétiques stationnaires de ce genre [3, 4, 9, 13, 14, 18, 54, 56]. Ces exemples sont uniques en ce sens qu'ils décrivent des « coupes génétiques » à travers les populations de différentes espèces, coupes pratiquées au cours d'intervalles de temps commensurables, dans certains cas, avec la durée d'existence de la population dans le milieu concret.

Ces dernières années, nous avons tenté de simuler ces phénomènes dans des expériences avec *Drosophila melanogaster*. Deux types de structure de populations ont été utilisés, l'un rappelait le modèle insulaire de Wright, l'autre correspondait au modèle échelonné circulaire unidimensionnel de Kimura [7, 8, 10, 12]. Dans le premier cas, on suivait la dynamique des fréquences d'allèles pour le locus autosomique estérase-6. Dans le second, on étudiait, outre l'estérase, la dynamique des fréquences d'allèles pour le locus

α-phosphoglycérate-déshydrogénase. Dans les deux modèles, les générations n'étaient pas superposables avec cette seule différence que dans le modèle insulaire, les cycles à migration entre les composantes du système alternaient avec des cycles à isolement, alors que dans le modèle échelonné, la migration était constante, son intensité et sa direction ne variaient pas durant toute l'expérience. Pratiquement, nous sommes parvenus à reproduire exactement les particularités de la génétique des populations subdivisées étudiées

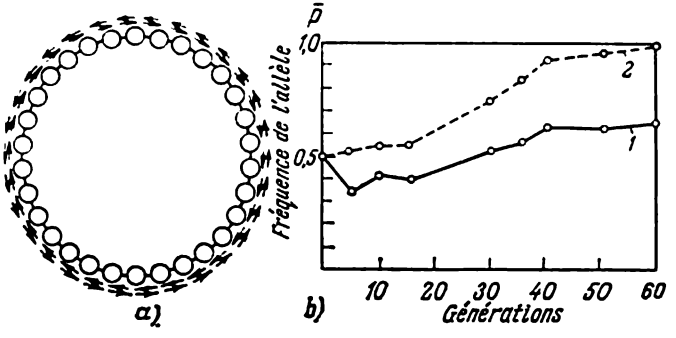


Fig. 4. Particularités du processus génétique dans une population subdivisée expérimentale:

a — modèle échelonné circulaire de population de *Drosophila melanogaster* composée de 30 subpopulations; l'effectif moyen du système est de 4000 mouches, le coefficient de migration est ~0,03. Les flèches indiquent le sens de la migration; *b* — dynamique de la fréquence de l'allèle F de l'α-phosphoglycérate-déshydrogénase dans un système de populations expérimental (*1*) et dans une population panmictique (*2*) ayant un patrimoine génétique identique

auparavant dans la nature et à montrer ainsi la stabilité de la composition génétique du système de populations, même en ce qui concerne les fréquences d'allèles du locus soumis à une puissante sélection orientée (fig. 4). Dans les deux séries de l'expérience, le système de populations conservait, au fil des générations, la caractéristique génétique de la population ancestrale, bien qu'en l'absence du noyau du système, ces fréquences puissent n'être nullement propres aux populations actuellement vivantes et ne se reconstruire que si l'on effectue la moyenne sur toutes les composantes structurales. Ainsi, la divergence génétique de l'espèce peut-elle être considérée comme une transformation successive du patrimoine génétique collectif de la population ancestrale au fur et à mesure de sa différenciation en subdivisions subordonnées au fil des générations et suivant l'aire d'habitation. Tout au moins, dans les conditions d'un environnement variant normalement, la fréquence génique moyenne et, une fois la phase de stabilité atteinte, la dispersion dans le système restent invariantes en ce qui concerne toutes les transformations microévolutives possibles, différemment orientées et, par consé-

quent, s'équilibrant mutuellement. En d'autres termes, une population isolée, si elle ne disparaît pas au cours de l'histoire, se déploie en elle-même pour parler métaphoriquement, entretenant un équilibre dynamique avec l'environnement. Le maintien de cette « mémoire » génétique dans le système est une condition nécessaire de sa stabilité et, en même temps, le témoignage de la conservation de la structure interne constituée historiquement.

Apparemment, plus l'organisation interne du système est compliquée, plus sa diversité interne est grande et plus elle résiste à différents facteurs extérieurs. De ce point de vue, la plus grande stabilité spatiotemporelle doit être propre à une espèce largement distribuée. Il est difficile de se représenter comment des processus déterminant la stabilisation maximale de l'espèce en tant que système de populations unique pourraient en même temps être à l'origine d'espèces nouvelles. Pour répondre à cette question, l'approche traditionnelle de la génétique des populations ne suffit pas. Il faut s'adresser à la dernière étape du développement de la génétique, directement liée aux aspects moléculaires de l'organisation du génome des eucaryotes, et à l'examen de la variabilité des protéines en tant que marqueurs génétiques dans les populations de différentes espèces.

Pour comprendre les lois de la formation des espèces, il ne suffit pas de se borner au phénomène du polymorphisme héréditaire, il faut en même temps étudier le système génétique monomorphe de l'espèce

Le polymorphisme héréditaire des protéines — clé de l'étude des phénomènes génétiques dans les populations — attire encore l'intérêt attentif de nombreux généticiens pour d'autres raisons. La plus importante est l'ampleur extrême de la variabilité génétique biochimique mise en évidence, au cours des 10 à 15 dernières années, dans les populations d'espèces les plus diverses. Lewontin et Hubby sur la drosophile [36, 42], Harris sur l'homme [34], Manwell et Baker sur les oiseaux [43] ont montré que, dans de nombreuses populations naturelles, le polymorphisme est découvert jusqu'à 30 % et plus parmi les loci étudiés et qu'un individu moyen peut être hétérozygote pour 12 % de ces gènes. Plus tard, ces découvertes se virent confirmées sur d'autres espèces [50].

Si l'on admet que la drosophile et l'homme ont au moins 10 000 à 50 000 gènes actifs, ceci signifie que dans une certaine population moyenne pas moins de 3 000 à 15 000 d'entre eux doivent être représentés par des allèles et qu'un individu moyen peut être hétérozygote pour plusieurs centaines et même milliers de ces gènes.

Le nombre des travaux dans ce domaine croît continuellement, faisant renaître, à un autre niveau d'analyse, la vieille discussion

pour savoir lequel des facteurs de la dynamique des populations — la *dérive génétique* ou la *sélection* — est responsable d'une diversité génétique inouïe.

Actuellement, se sont nettement dessinés deux points de vue polaires dont l'un a tendance à considérer le polymorphisme des protéines comme sélectivement neutre (M. Kimura et al.), et l'autre, comme une variabilité d'adaptation entretenue par diverses formes de sélection (F. Ayala et al.).

R. Lewontin, dont les travaux pionniers de 1963-1965 ont contribué pour beaucoup à développer cette direction des recherches, occupe une position intermédiaire soulignant un certain nombre de difficultés rencontrées dans la voie de l'adoption de tel ou tel système d'idées. Cette position se voit nettement dans son livre *The genetic basis of evolutionary changes* [40].

Il est évident que si le polymorphisme des protéines est stable et est entretenu par la prédominance sélective des hétérozygotes, dans ce cas (avec un modèle multiplicatif d'adaptabilité) les populations naturelles doivent porter un énorme fardeau génétique (fardeau de ségrégation):

$$\tilde{L} = \sqrt{\sum_{1}^{n} \left[\frac{S_1 S_2}{S_1 + S_2} \right]^2 2 \log_e (0{,}4N)}$$

où S_1 et S_2 sont des coefficients de sélection dans le groupe considéré de loci superdominants dialléliques (n),

N est l'effectif de la population.

Par exemple, si $S_1 = S_2 = 0{,}1$, $n = 1000$ et $N = 25\,000$, $\tilde{L} = \sqrt{46} = 6{,}8$. Ceci signifie que pour que l'effectif de la population reste stable au fil des générations, l'individu le plus adapté de cette population doit produire 898 fois plus de descendants en comparaison du phénotype moyen! Ces calculs représentent un argument de poids en faveur des vues de M. Kimura qui font écho pour beaucoup aux idées de H. Muller selon lesquelles la plupart des mutations nouvelles sont nocives pour l'organisme. Donc, si dans les populations naturelles il existe une variabilité génétique des protéines aussi gigantesque, elle restera neutre pour la sélection (« école néoclassique » selon la terminologie de R. Lewontin).

D'autre part, on a obtenu de nos jours assez de preuves de l'avantage sélectif des individus hétérozygotes pour telle ou telle variante électrophorétique de protéines comme cela fut démontré, par exemple, dans la partie précédente de l'article (v. p. 13). Naturellement, les données de ce genre viennent confirmer les vues des « sélectionnistes » comme R. Lewontin appelle les tenants de l'« hypothèse de l'équilibre ».

L'auteur de cet article n'a pas pour but d'analyser en détail l'argumentation de fond des deux parties de la polémique. Il n'examinera pas non plus dans quelle mesure sont fiablement estimés les paramètres de la structure de populations utilisés en calcul du fardeau génétique et, avant tout, un paramètre si essentiel pour toute recherche en génétique des populations : la partie génétiquement active de la population.

L'auteur indiquera seulement que, dès 1968-1972 [1, 2, 4, 15], une hypothèse a été émise qui, apparemment, expliquait de façon satisfaisante la variété génétique des populations naturelles d'après les diverses protéines à fonction enzymatique ou non. Notre approche réunit les vues des écoles « de l'équilibre » et « néoclassique » et postule la dualité dans l'organisation structuro-fonctionnelle du génome des eucaryotes. Dans ce système d'opinions, on considère le polymorphisme des protéines comme une variabilité relativement neutre, liée aux propriétés adaptatives secondaires de l'espèce, et les protéines génétiquement monomorphes comme des marqueurs de fonctions primordiales dont la variabilité est biologiquement inadmissible. Étant nocives, les mutations dans cette partie du génome doivent être immédiatement éliminées par la sélection, surtout aux stades ontogénétiques précoces.

La phénoménologie du polymorphisme biochimique, la réalité du phénomène même ne soulèvent aucun doute. En ce qui concerne le monomorphisme génétique, étant donné qu'il ne permet pas d'étudier la divergence génétique de l'espèce, il reste généralement en dehors du champ optique des chercheurs. Il n'est pas exclu que cette invariabilité puisse simplement refléter le volume insuffisant d'un échantillon, la propriété d'une population particulière ou la conséquence d'un pouvoir de discrimination insuffisant de la méthode. Enfin, la thèse d'une liaison absolue de la protéine génétiquement monomorphe avec la viabilité de l'organisme demande encore à être démontrée.

Grâce aux travaux effectués tant par notre groupe que par d'autres chercheurs, ces objections peuvent maintenant être réfutées.

1° On peut considérer comme un fait établi que l'invariabilité génétique enregistrée lors de l'électrophorèse des protéines ne peut être attribuée dans tous les cas au pouvoir de discrimination insuffisant de la méthode. On sait depuis longtemps que dans l'examen complet de la structure primaire des enzymes, leur centre catalytique s'avère souvent inchangé, alors que des secteurs fonctionnellement moins importants de la molécule se modifient. Par exemple, chez diverses espèces de vertébrés la lactate-déshydrogénase a une composition différente en acides aminés. Cependant, le centre catalytique de cette enzyme constitué de 12 radicaux est invariable dans tous les cas étudiés. Les radicaux d'histidine localisés sur les sites 58, 87, 63 et 92 des chaînes α et β de la molécule de globine et qui sont fonc-

tionnellement les plus importants restent tout à fait inchangés. Pour de nombreux groupes de vertébrés, ce monomorphisme se maintient depuis 500 millions d'années de leur histoire naturelle [37]. La substitution de la tyrosine à n'importe laquelle de ces histidines provoque une hémoglobinopathie chez l'homme. Il serait aisé d'augmenter le nombre d'exemples de ce genre [21, 38].

En outre, la réalité de l'état monomorphe de certains loci protéiques a reçu ces derniers temps une confirmation malgré la découverte des allèles dits thermiques.

R. Lewontin a montré, dans son rapport au XIVe Congrès international de génétique, que le perfectionnement ultérieur des méthodes de mise en évidence des variations héréditaires au niveau moléculaire ne change rien à l'idée de l'existence de deux groupes de loci responsables de la synthèse des protéines. Selon lui, on peut considérer comme une règle que les loci qui s'avèrent polymorphes dans un certain procédé d'investigation, font augmenter la diversité des allèles si l'on utilise des méthodes plus perfectionnées, alors que les loci qui apparaissent comme monomorphes à l'électrophorèse ordinaire, restent monomorphes avec une investigation plus poussée [41].

En 1979, Brown et Langley [22] ayant étudié au moyen de l'électrophorèse bidimensionnelle certaines protéines tissulaires de la drosophile, avaient découvert une variabilité génétique si faible qu'ils ont conclu à réviser les estimations faites antérieurement des niveaux de polymorphisme et d'hétérozygotie des populations naturelles. Cependant, il faut rappeler qu'un résultat très proche avait déjà été obtenu en 1971 [16] en incluant dans l'exploration électrophorétique plusieurs groupes de protéines hydrosolubles à fonction non enzymatique.

2° On peut estimer prouvé que, comme il était postulé, le monomorphisme génétique représente non seulement un phénomène caractérisant la population, mais aussi un phénomène caractérisant l'espèce dans son ensemble. Ce fait est établi d'une façon assez fiable pour *Homo sapiens*, dont les représentants de toutes les races principales sont aujourd'hui étudiés d'après un nombre suffisant de marqueurs protéiques des gènes, parmi lesquels certains sont monomorphes sur toute l'aire d'habitation de l'espèce [49]. Compte tenu des données de nos collègues américains, canadiens et japonais, notre groupe a réussi à obtenir des preuves du monomorphisme de certains systèmes protéiques des salmonidés du Pacifique. Ces preuves reposent sur l'analyse électrophorétique de plusieurs milliers de spécimens recueillis sur de vastes espaces du littoral asiatique et américain du Pacifique Nord [1, 4, 6, 15, 16, 39, 52].

3° On a obtenu des preuves directes du lien des systèmes protéiques monomorphes avec la viabilité de l'organisme en étudiant des groupes d'enfants ayant des malformations multiples et grossières

[31, 32]. On a montré que si la fréquence moyenne pour le locus par individu de variantes rares de protéines était de $3{,}5 \cdot 10^{-4}$ pour le groupe d'enfants témoins, elle était de l'ordre de $2{,}7 \cdot 10^{-3}$ pour le groupe d'enfants ayant de nombreuses malformations grossières. Ceci signifie qu'environ 10 % du nombre total de mutations géniques survenant dans une génération donnée parviennent à la partie reproductrice de la population et sont transmises aux générations successives comme plus ou moins neutres. En effet, ces estimations coïncident parfaitement avec la valeur du coefficient de sélection négative, qui est d'environ 90 % pour le groupe d'enfants étudiés par nous. Compte tenu des données nouvelles, on révise en même temps les évaluations des cadences de mutation pour les loci protéiques chez l'homme, établies jusqu'à présent par des méthodes indirectes [35, 45, 46, 48].

L'obtention de preuves directes de l'importance sélective négative de rares variantes électrophorétiques de protéines permet une explication nouvelle de la redondance, dans les populations, de divers allèles électrophorétiques rares lors de l'approximation des distributions empiriques de leurs fréquences par la fonction mathématique correspondante supposant la neutralité du polymorphisme protéique. A notre laboratoire, Y. Doubrovà [33] a effectué l'analyse requise. Bien que le travail ne soit pas achevé, on peut déjà dire que l'approximation s'avère tout à fait satisfaisante en adoptant notre modèle selon lequel l'existence d'allèles rares dans la population reflète dans une grande mesure l'équilibre entre la sélection stabilisante et la mutation spontanée.

Donc, nous avons obtenu, ces derniers temps, des données qui non seulement confirment l'idée du monomorphisme génétique de l'espèce en tant que phénomène naturel, mais aussi versent la lumière sur les mécanismes entretenant la stabilité du phénotype au niveau protéique. Parmi elles, les plus importantes sont les suivantes.

1° **Pression de sélection stabilisante.** Si la protéine a une structure quaternaire et (ou) si elle est responsable des fonctions les plus importantes, la grande majorité des mutations seront interdites, c'est-à-dire qu'elles seront éliminées par la sélection aux stades les plus précoces de l'ontogénèse. Ce sont les mutations intéressant le centre actif de l'enzyme ou les parties de la molécule qui sont indispensables à l'assemblage des sous-unités.

2° **Particularités de l'organisation du matériel génétique.** S'il y a dans la synthèse protéique un mécanisme de correction ou d'élimination des mutations, son rôle majeur dans le maintien de la structure primaire de la protéine est évident. On aurait pu s'y attendre même si l'hypothèse master-slave de Callan était juste. Cependant, depuis la découverte des gènes « mosaïques » [25, 30], une possibilité de ce genre devient plus évidente si l'on considère les introns comme une sorte de piège pour les mutations.

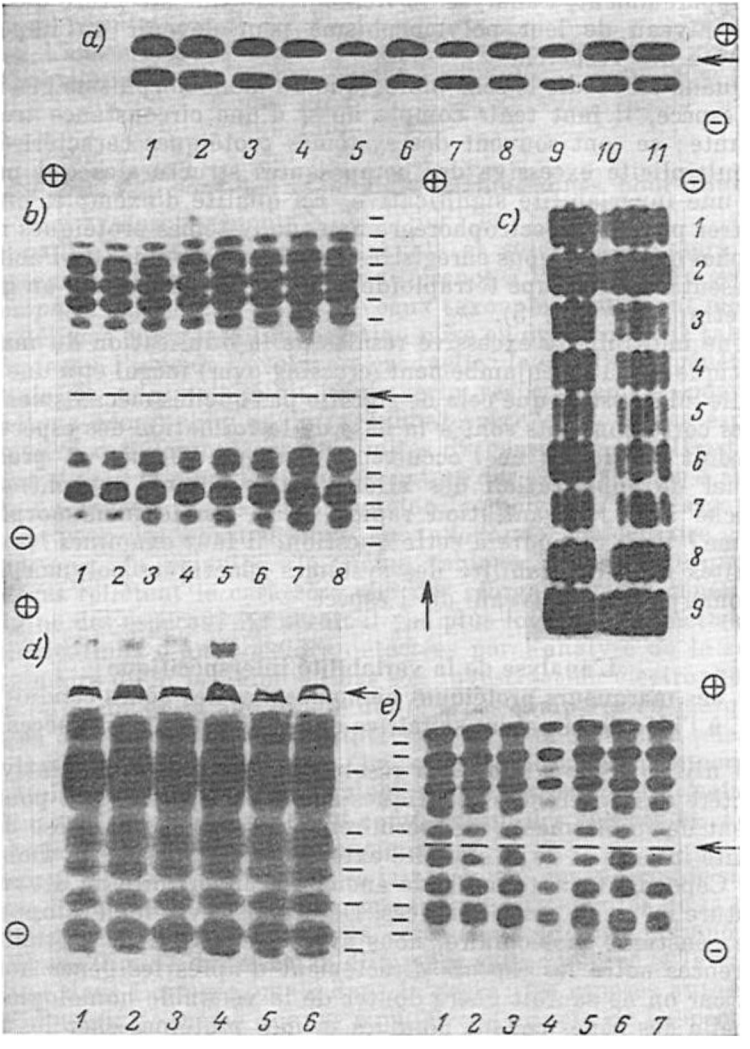


Fig. 5. Formes moléculaires multiples des protéines chez les salmonidés du Pacifique:

a — hémoglobines de 11 exemplaires de saumon kéta (*Oncorhynchus keta* Walb.), électrophorèse en gélose, tampon au barbital, pH 8,6; *b* — hémoglobine de 8 exemplaires de saumon kéta, électrophorèse en gel d'amidon, tampon au borate, pH 8,5; *c* et *e*, la même chose pour quelques exemplaires de saumon gorbuscha (*O. gorbuscha* Walb.); cette pluralité des hémoglobines des salmonidés est due à la présence chez eux de 7 ou 8 subdivisions diverses formant des molécules tétramères [58, 59]; *d* — protéines hydrosolubles du cristallin de l'œil du saumon kéta, électrophorèse sur gélose.

Sur toutes les photos on voit l'absence de variabilité individuelle dans le nombre ou la position des divers composantes protéiques: monomorphisme [4]

Apparemment, l'analyse du lien entre la fine structure des gènes et le niveau de leur polymorphisme peut devenir un important domaine des recherches.

Quand on étudie le sens biologique du monomorphisme génétique de l'espèce, il faut tenir compte aussi d'une circonstance très importante : ce sont souvent des systèmes protéiques caractérisés par la multiplicité excessive des composantes structurales qui présentent une invariabilité significative. En qualité d'exemple, on peut montrer plusieurs électrophorégrammes de systèmes protéiques monomorphes que nous avons enregistrés chez les salmonidés du Pacifique, représentant un groupe tétraploïde ou, très probablement, un groupe allotétraploïde (fig. 5).

Une multiplicité excessive résulte de la duplication du matériel génétique due à un enjambement (crossing-over) inégal et à une polyploïdie. Mais est-ce que cela ne signifie pas que les mécanismes génétiques correspondants sont à la base de la formation des espèces que l'on doit considérer en l'occurrence non pas comme un processus graduel de substitution des allèles, mais comme un phénomène rattaché à la réorganisation rapide de la partie monomorphe du génome ? Pour répondre à cette question, il faut examiner les particularités de la variabilité des systèmes génétiques polymorphes et monomorphes au niveau de l'espèce.

L'analyse de la variabilité interspécifique des marqueurs protéiques monomorphes des gènes conduit à l'idée de la nature saltatrice de la formation des espèces

Il n'y a pas longtemps, la résolution des questions relatives au caractère des différences génétiques entre les espèces avait pour fondement des croisements qui permirent de confirmer la thèse de « la fertilité interne et de la stérilité externe de l'espèce » (Th. Dobzhansky). Cependant, cette méthode ancienne et méritoire ne dit rien de la nature même de ces différences. Désormais, avec le développement de la génétique moléculaire, nous avons la possibilité d'étudier les différences entre les espèces directement d'après les gènes homologues, car on ne saurait guère douter de la véritable homologie fonctionnelle des gènes codant pour les mêmes protéines chez les représentants d'espèces non seulement rapprochés taxonomiquement, mais aussi éloignés. Jusqu'à présent, des confrontations de ce genre se faisaient en utilisant différentes mesures des distances génétiques ou des indices de similitude sans diviser les loci en poly et monomorphes. Selon un de ces procédés proposés par Nei [47], l'identité normalisée des gènes entre deux populations pour le locus j

$$I_j = \frac{\sum X_i Y_i}{(\sum X_i^2 \sum Y_i^2)^{1/2}}$$

où X_i et Y_i sont respectivement les fréquences de l'allèle i dans les populations X et Y.

Pour l'ensemble des loci étudiés, l'identité génétique générale des échantillons

$$I = \frac{J_{xy}}{(J_x J_y)^{1/2}}$$

où J_x, J_y et J_{xy} sont les moyennes arithmétiques pour tous les loci, respectivement $\sum X_i^2$, $\sum Y_i^2$ et $\sum X_i \sum Y_i$.

En utilisant cette mesure, F. Ayala et ses collaborateurs [20] ont entrepris l'étude des degrés de différence à l'intérieur du groupe de *Drosophila willistoni* à divers niveaux taxonomiques et ont montré le parallélisme entre la différenciation mise en évidence et le tableau attendu face à l'idée de l'évolution graduelle : la similitude génétique maximale et minimale était découverte respectivement pour les populations géographiques de la même espèce et pour des espèces notoirement différentes dans les limites du genre (fig. 6). Il est évident que si nous faisions l'étude de la nature des substitutions d'acides aminés au niveau d'une seule protéine, comme cela est largement pratiqué pour les constructions en taxonomie moléculaire, nous obtiendrions, en principe, le même tableau. Mais est-ce que ces différences reflètent le caractère réel des réorganisations génétiques à l'origine des espèces ? Ne serait-il pas plus logique d'admettre que les substitutions d'aminoacides détectées par l'analyse de la structure primaire de la protéine ou par son investigation électrophorétique ne reflètent que l'ancienneté des groupes comparés en marquant le temps de leur séparation évolutive et de leur existence relativement indépendante ultérieure ? On comprend que les groupements intraspécifiques plus jeunes devaient accumuler beaucoup moins de substitutions d'acides aminés en comparaison des espèces et, à plus forte raison, des genres plus anciens.

Pour se représenter comment se fait le calcul des indices de similitude ou d'identité, il suffit d'examiner les particularités de la différenciation interspécifique d'après certaines protéines polymorphes homologues (fig. 7). En effet, la différence entre l'espèce et la variété n'est pas dans l'essence, mais dans le degré : les espèces autonomes se distinguent, de même que les populations au sein de l'espèce, par la fréquence de rencontre des mêmes allèles ou par les gènes particuliers propres à elles seules. Evidemment, étant donné que les espèces sont génétiquement fermées, ce dernier genre de différence se rencontrera au niveau interspécifique et non pas au niveau intraspécifique en accord avec la plus grande échelle des différences.

Etudions maintenant le caractère des différences entre des espèces notoirement diverses d'après les protéines génétiquement monomorphes. Dans ce but, adressons-nous aux résultats de l'exploration électrophorétique des hémoglobines des salmonidés du Pacifique

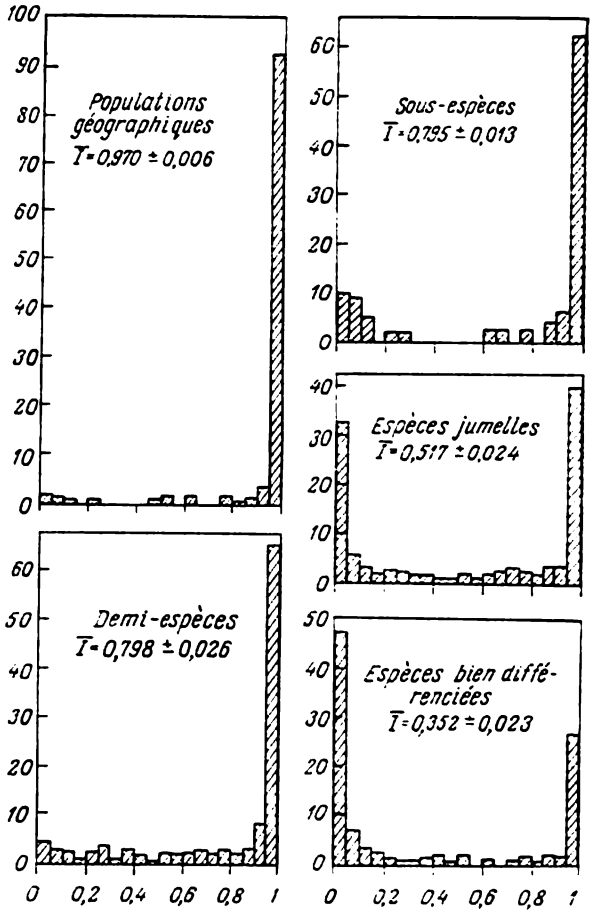


Fig. 6. Distribution des loci d'après l'identité des fréquences alléliques aux différents stades de la divergence évolutive dans le groupe des formes de *Drosophila willistoni* [20].

ainsi qu'aux données relatives à quelques groupes de protéines génétiquement monomorphes chez d'autres espèces animales (fig. 8 et 9). Dans tous les cas étudiés, des différences qualitatives sont suivies : d'après ces caractères monomorphes les espèces sont aussi discontinues et uniques que sont différents les génotypes pour tel ou tel système de polymorphisme génétique au sein de l'espèce. En d'autres termes, *selon la partie génétiquement monomorphe du génome toute espèce se présente comme un individu distinct.* Bien plus, quand on parvient à comparer aux espèces parentes des hybrides interspécifi-

ques rares ou des espèces d'origine hybride, on découvre une dépendance très importante: les caractères monomorphes spécifiques se conduisent comme des unités génétiques entières représentant une simple somme des types fondateurs, la formation de produits hybrides ou même des rapports de prédominance-récessivité (v. fig. 9). C'est ainsi que se manifeste la partie monomorphe du

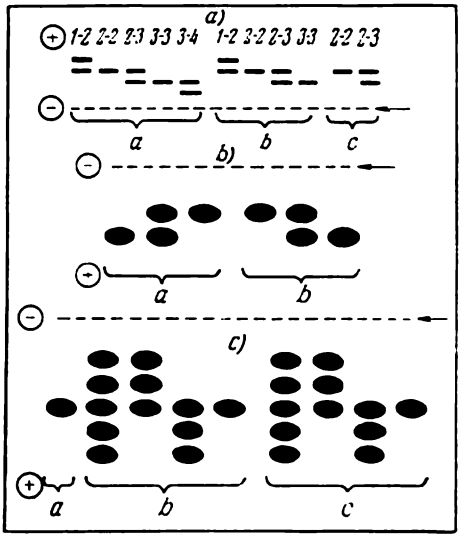


Fig. 7. Variabilité interspécifique des protéines génétiquement polymorphes: *a*) — génotypes pour le locus estérase sérique chez trois espèces de thons: *a* — *Thunnus maccoyii*, *b* — *T. obesus*, *c* — *T. albacares* (électrophorèse sur gel d'amidon); *b*) — génotypes de l'albumine sérique des sébastes: *a* — *Sebastes marinus*, *b* — *S. mentella* (électrophorèse sur gélose); c) — génotypes de la malate-déshydrogénase musculaire chez trois espèces de salmonidés du genre *Oncorhynchus*: *a* — *O. nerka*, *b* — *O. gorbuscha*, *c* — *O. keta* (électrophorèse sur gel de polyacrylamide [4]

génome marquée par des familles isofonctionnelles de gènes multiples et protégée d'une certaine façon de la ségrégation.

Il est évident que si nous voulions déterminer la similitude génétique suivant les systèmes protéiques monomorphes codés par de multiples gènes et différenciant qualitativement les espèces notoirement diverses, nous obtiendrions la valeur 1 pour n'importe quelles populations au sein de l'espèce et des valeurs croissantes des différences aux niveaux spécifiques ou taxonomiques plus élevés. Or, il est clair que ces valeurs ne refléteraient pas un passage graduel de la population à l'espèce, mais une rupture qualitative ou, en d'autres termes, un saut.

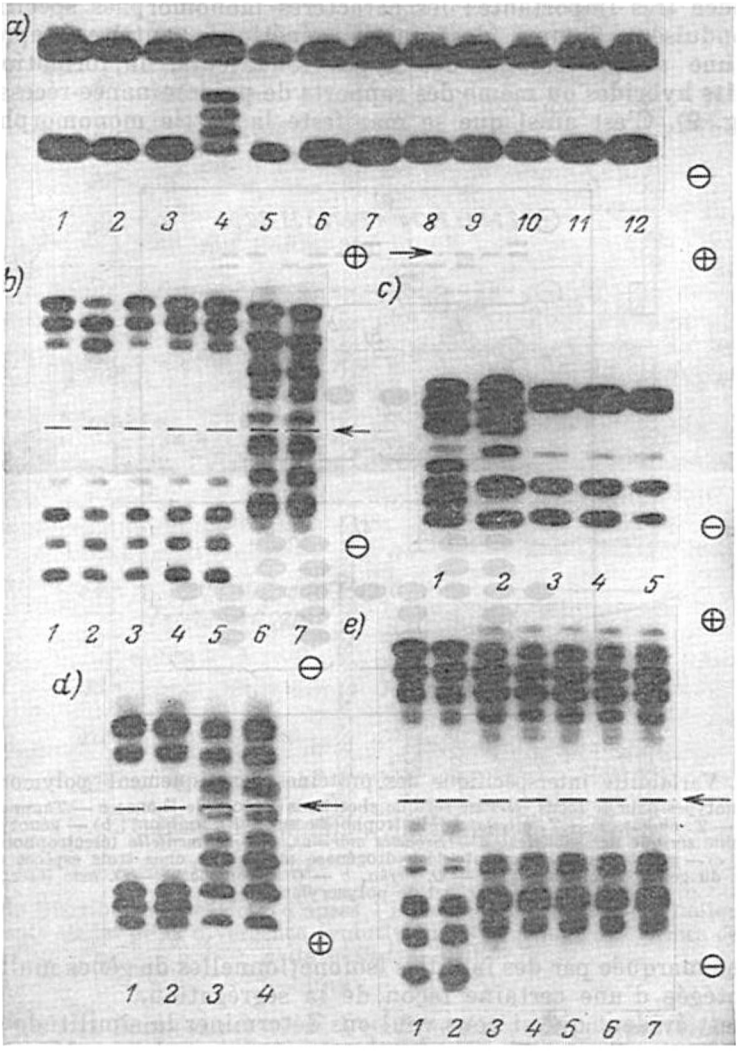


Fig. 8. Spécificité d'espèce découverte par électrophorèse de multiples hémoglobines de salmonidés:

a — 1-3 et 5-12 — Oncorhynchus keta, 4 — O. kisutch (électrophorèse sur gélose); *b — 1-5 — Salvelinus leucomaenis, 6-7 — O. gorbuscha* (électrophorèse en gel d'amidon); *c — 1 — O. gorbuscha, 2* — hybride (F_1), *3-5 — O. masu* (électrophorèse sur gélose); *d — 1-2 — O. keta, 3-4 — O. gorbuscha*; *e — 1-2 — Salvelinus leucomaenis, 3-7 — O. keta* (électrophorèse en gel d'amidon) [4]

Si discréditée que soit l'idée de la formation des espèces par sauts en raison de conceptions dogmatiques en biologie, on ne peut comprendre objectivement la microévolution sans tenir compte de tout

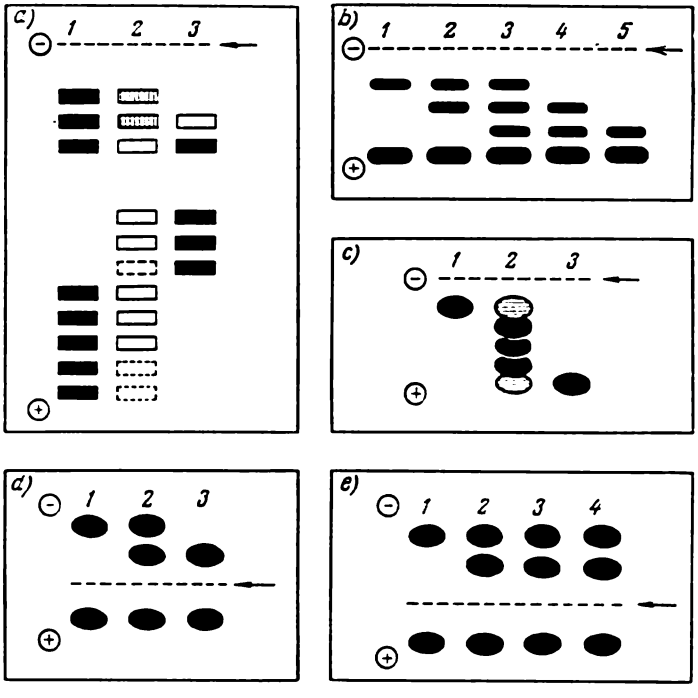


Fig. 9. Spectres électrophorétiques de protéines spécifiques d'espèce chez des hybrides interspécifiques en comparaison des espèces originelles:

a — peroxydase des feuilles: *1* — de *Nicotiana sylvestris*, *2* — de *N. tomentosiformis*, *3* — de l'hybride interspécifique F_1 (électrophorèse sur gel de polyacrylamide); *b* — protéines musculaires hydrosolubles des poissons: *1* — *Richardsonius balteatus*, *5* — *Mylocheilus caurinum*, *3* — des hybrides F_1, *2* et *4* — des back cross correspondants (électrophorèse sur gel d'amidon); *c* — lactate-déshydrogénase du myocarde des lézards du genre *Cnemidophorus*: *1* — *C. gularis*, espèce bissexuée, *3* — *C. tigris*, espèce bissexuée, *2* — *C. tesselatus*, espèce parthénogénétique d'origine hybride (électrophorèse en gel d'amidon); *d* — hémoglobines des oiseaux: *1* — *Coturnix coturnix*, *3* — *Gallus gallus*, *2* — de l'hybride correspondant F_1 (électrophorèse sur gel d'amidon); *e* — chaînes globiniques de l'hémoglobine: *1* — de l'âne, *2* — du mulet, *3* — du bardot, *4* — du cheval (électrophorèse en gel d'amidon). Les rayures différentes reflètent la variation d'intensité des fractions sur les électrophorégrammes originaux [15]

l'ensemble de faits et d'approches connus dans la génétique biochimique moderne des populations.

Et si l'on conclut que le monomorphisme génétique de l'espèce est un phénomène naturel aussi réel que son polymorphisme, la seule reconnaissance de ce fait, compte tenu du caractère de la variabilité

interspécifique des propriétés génétiquement monomorphes codées par des gènes multiples, offre la possibilité de traiter la formation des espèces non pas comme un processus probabiliste graduel se déroulant au niveau des populations, mais comme la conséquence de remaniements qualitatifs de la partie monomorphe du génome. Etant donné que ce sont généralement les protéines multiples codées par des gènes dupliqués multiples qui sont les plus monomorphes et, en même temps, les plus chargées de propriétés spécifiques de l'espèce, les mécanismes connus des duplications géniques doivent refléter les réorganisations génétiques qui, s'avérant « compatibles » avec l'ontogénèse, peuvent, en une ou plusieurs démarches ultérieures, conduire un individu à l'isolement reproductif.

Les duplications du matériel génétique ayant pour base une autocopie locale excédentaire des gènes, la polyploïdie ou un enjambement (crossing-over) inégal, traduisent d'autres remaniements qualitatifs du génome que la mutation ponctuelle à l'origine du polymorphisme héréditaire des protéines. Au fond, les translocations, les délétions et les fusions des chromosomes entraînent les mêmes changements d'ordre des gènes dans le génome et les variations correspondantes de l'expression génique [19].

L'importance de tels remaniements du matériel génétique pour l'évolution est aujourd'hui prouvée pour de nombreux groupes taxonomiques, surtout pour ceux auxquels on donnait récemment encore une place particulière pour faire ressortir les différences qualitatives des phénomènes évolutifs dans le règne animal et végétal. Le sens biologique essentiel de ces remaniements consiste en ce qu'ils transposent par saut tous les gènes dans le génome ou leur lot déterminé dans un état hétérozygote constant et, par conséquent, assurent aux individus les avantages d'un niveau d'adaptation qualitativement différent, en débarrassant la population future du fardeau de ségrégation. En même temps, ceci signifie une augmentation de la fiabilité du stockage et de la transmission aux générations suivantes de l'information concernant les fonctions vitales qui reflètent l'unicité de l'espèce nouvelle.

Dans ce schéma de la microévolution par remaniement de la partie monomorphe du génome, il existe un endroit faible. Plus haut, nous avons souligné que ce genre de variabilité se trouve sous le contrôle vigilant de la sélection stabilisante. Comment donc, dans ce cas, la formation des espèces peut-elle se réaliser ?

Pour résoudre cette question, il est nécessaire de postuler un état particulier de l'environnement dans lequel une macromutation peut s'avérer neutre ou bénéficier de certains avantages sélectifs.

Nous supposons que les réorganisations correspondantes du génome ont constamment lieu au sein des espèces diverses, mais elles n'acquièrent le droit d'exister que grâce aux variations du milieu naturel suivies de la rupture des barrières historiquement constituées

entre les populations (et entre les espèces) qui élève brusquement les niveaux de polymorphisme génétique des populations et d'hétérozygotie des individus. Le polymorphisme engendre un polymorphisme encore plus grand [51], le volume du fardeau génétique augmente fortement et, par conséquent, l'espèce meurt ou se transforme en une autre espèce. De plus, grâce au passage des systèmes du polymorphisme génétique à l'état monomorphe, la variabilité habituelle s'éteint. Dans quelle mesure de telles réorganisations du matériel génétique sont aléatoires et dans quelle mesure peuvent-

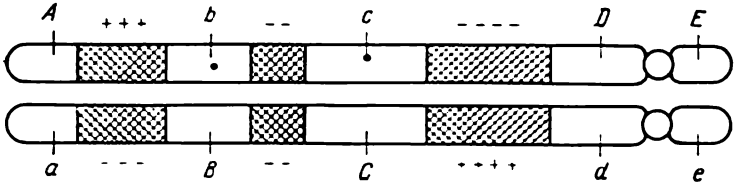


Fig. 10. Modèles de systèmes génétiques « fermé » et « ouvert »:
Paire d'homologues hétérozygotes pour 12 loci sur 14. Par enjambement (crossing-over) à l'intérieur du système « ouvert » (non rayé), il se produit parfois des zygotes assez bien adaptables. Dans les limites du système « fermé » (rayé), la plupart des cross-over ne sont pas adaptables.
Les lettres désignent les gènes ou polygènes communs. Les plus et les moins représentent des complexes intérieurement équilibrés parmi lesquels n'importe quelles combinaisons « plus-moins » dans les limites d'un bloc quelconque sont faiblement adaptées. L équilibre génétique protège la plupart des blocs contre la fixation en état homozygote [23]

elles être traitées comme le développement du taxon déterminé par l'histoire précédente, c'est une question qui reste ouverte.

Dans son article *The genetics of speciation at the diploid level* publié dans *American Naturalist*, H. Carson [23] développe des conceptions, proches des nôtres, de la formation des espèces. Selon Carson, chez les espèces bisexuées il faut distinguer deux systèmes de variabilité génétique: « ouvert » et « fermé ». Le premier est représenté par des loci à ségrégation libre, n'influant pas notablement sur la viabilité (polymorphisme de différents types, variabilité subspécifique, clinale); le second, par des blocs de gènes (supergènes) intérieurement équilibrés, coadaptés et influant si fortement sur l'adaptabilité que n'importe laquelle de leurs réorganisations est éliminée par la sélection naturelle (fig. 10). Ces blocs protégés de l'enjambement (crossing-over) par l'inversion ou simplement par de fortes interactions épistatiques des gènes, varient entre les espèces, mais non pas au sein de l'espèce. La formation des espèces a pour base un remaniement du système fermé du génome et les nouvelles espèces naissent d'un seul ou d'un petit nombre d'individus.

Le parallélisme des vues de H. Carson et des nôtres est évident. La différence ne réside que dans la terminologie, les méthodes de

recherche, les objets et certains autres moments rattachés à l'interprétation de l'état de l'environnement lors de la formation des espèces (fig. 11). Il faut indiquer seulement que, ces idées étant sous telle ou telle forme critiquées ou réfutées par certains biologistes, J. Powell [53] aux U.S.A. a réalisé, au cours des cinq dernières années, une vaste expérience de simulation de ce schéma de formation des espèces et a le premier obtenu, en génétique des populations expé-

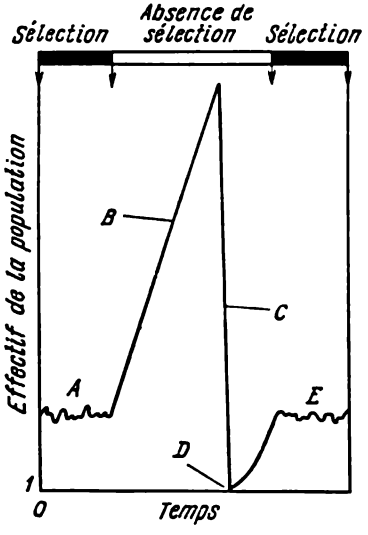


Fig. 11. Schéma de la formation des espèces par remaniement du système génétique « fermé »:

Après la phase de sélection (*A*), la population passe par le stade d'épanouissement (*B*) et de catastrophe (*C*). Les individus fondateurs (*D*) ayant survécu à la catastrophe produisent une nouvelle population, stabilisée de nouveau par la sélection naturelle (*E*).
Les remaniements du système génétique « fermé » de l'espèce « compatibles » avec l'ontogénèse ne sont possibles que dans la phase d'affaiblissement de la sélection où un ou plusieurs individus « discordants » peuvent survivre [23]

rimentale, sur 50 ou 60 générations seulement un isolement reproducteur stable entre des populations qui, initialement, s'entrecroisaient librement.

* * *

L'analyse effectuée a montré que les thèses exposées au début de l'article reçoivent une confirmation complémentaire grâce aux élaborations nouvelles dans le domaine de la génétique biochimique et expérimentale des populations. Il est évident que les données obtenues et le modèle de formation des espèces construit sur leur base ont un intérêt non seulement cognitif mais aussi pratique, surtout en ce qui concerne l'appréciation du caractère des processus génétiques dans les populations humaines actuelles et le pronostic de notre rôle possible dans les phénomènes évolutifs de l'avenir. S'il existe des systèmes génétiquement monomorphes (ou fermés selon la terminologie de Carson) de caractère spécifique général, l'existence de toute espèce sera liée à la stabilité et à l'intégrité de ces

systèmes mêmes qui, dans la norme, doivent être fermés à la destruction par une pression de mutation ou de recombinaison génique. Il est évident que les conditions d'accélération de la cadence des mutations ou de croissance de l'intensité de tel ou tel facteur entraînant la destruction de ces systèmes, peuvent impliquer des conséquences funestes d'une grande portée, aboutissant soit à l'élimination de l'espèce, soit à sa transformation en une autre espèce. Toutes ces conceptions, pour beaucoup déductives, acquièrent un retentissement particulier de nos jours où l'humanité vit une phase qualitativement nouvelle de son histoire, marquée par la rupture, sous nos yeux, des barrières migratoires, la détérioration de la structure de populations, la croissance de l'intensité de la recombinaison et, probablement, de la mutation des gènes.

Du point de vue de la théorie généralement admise de l'évolution, toute transformation évolutive doit être considérée comme un phénomène progressif. Or, dans quelle mesure peut-on, en faveur de ce progrès, renoncer au legs social et génétique de nos ancêtres proches ou éloignés, on n'en sait rien. Il est possible que la réponse à cette question dépasse le cadre de la génétique évolutive et de la génétique des populations.

RÉFÉRENCES

[1] *Алтухов Ю. П.* О соотношении моно- и полиморфизма гемоглобинов в микроэволюции рыб. *ДАН СССР*, 1969, 189, 5, 1115-1117.

[2] *Altukhov Yu. P.* Biochemical polymorphism in fishes and the problem of intraspecific differentiation. In : Abstr. XIIth Internat. Conf. Animal Blood Groups and Biochem. Polymorphisms, 1970.

[3] *Altukhov Yu. P.* Rational management of marine biological resources in the light of population genetics. In : Propogation marine resources Pacific Ocean, Tokyo, Tokai Univ. Press, 1973, 31-39.

[4] *Алтухов Ю. П.* Популяционная генетика рыб. М., Пищ. пром-сть, 1974, 245 с.

[5] *Алтухов Ю. П.* Генетика природных популяций и ресурсы биосферы. *Вестн. АН СССР*, 1975, 10, 37-45.

[6] *Altukhov Yu. P.* The stock concept from the viewpoint of population genetics. *Canad. J. Fish. Aquat. Sci.*, 1981, 38, 1523-1528.

[7] *Алтухов Ю. П., Бернашевская А. Г.* Экспериментальное моделирование динамики генных частот в системе полуизолированных популяций. *ДАН СССР*, 1978, 238, 3, 712-714.

[8] *Алтухов Ю. П., Бернашевская А. Г., Милишников А. П.* и др. Экспериментальное моделирование генетических процессов в популяционной системе *Drosophila melanogaster*, соответствующей кольцевой ступенчатой модели. *Генетика*, 1979, 15, 4, 646-655.

[9] *Алтухов Ю. П., Калабушкин Б. А.* Стабильный полиморфизм в современной и ископаемой популяциях моллюска *Littorina squalida*. *ДАН СССР*, 1974, 215, 6, 1477-1480.

[10] *Алтухов Ю. П., Победоносцева Е. Ю.* Экспериментальное моделирование генетических процессов в подразделенных популяциях. *ДАН СССР*, 1978, 238, 2, 466-469.

[11] *Алтухов Ю. П., Победоносцева Е. Ю.* Исследование биологических особенностей экспериментальной популяционной системы *Drosophila melanogaster*. *Журн. общ. Биол.*, 1979а, 40, 3, 368-376.
[12] *Алтухов Ю. П., Победоносцева Е. Ю.* Особенности генетического процесса в экспериментальной популяционной системе *Drosophila melanogaster*. *Журн. общ. Биол.*, 1979б, 40, 6, 196-923.
[13] *Алтухов Ю. П., Пудовкин А. И., Салменкова Е. А., Коновалов С. М.* Стационарность распределений частот генов лактатдегидрогеназы и фосфоглюкомутазы в системе субпопуляций локального стада рыб. Сообщ. II. *Генетика*, 1975, 11, 4, 54-62.
[14] *Алтухов Ю. П., Рычков Ю. Г.* Популяционные системы и их структурные компоненты. Генетическая стабильность и изменчивость. *Журн. общ. Биол.*, 1970, 31, 5, 507-526.
[15] *Алтухов Ю. П., Рычков Ю. Г.* Генетический мономорфизм вида и его биологическое значение. *Журн. общ. Биол.*, 1972, 33, 281-300.
[16] *Алтухов Ю. П., Салменкова Е. А., Омельченко В. Т.* и др. О числе мономорфных и полиморфных локусов в популяции кеты — одного из тетраплоидных видов тихоокеанских лососей. *Генетика*, 1972, 8, 2, 251-259.
[17] *Алтухов Ю. П., Салменкова Е. А., Коновалов С. М.* и др. Стационарность распределения частот генов лактатдегидрогеназы и фосфоглюкомутазы в системе субпопуляций локального стада рыб (на примере *Oncorhynchus nerka* Walb.). Сообщ. I. *Генетика*, 1975, 11, 4, 44-53.
[18] *Алтухов Ю. П., Салменкова Е. А., Рябова Г. Д.* и др. Генетическая дифференциация популяций кеты и эффективность некоторых акклиматизационных мероприятий. *Биол. моря*, 1980, 3, 23-38.
[19] *Ayala F. J.* Molecular evolution. Sunderland (Mass.), Sinauer Assoc. Publ., 1976, 277 p.
[20] *Ayala F. J., Tracey M. L., Barr L. G.* and al. Genetic variation in natural populations of five *Drosophila* species and the hypothesis of the selective neutrality of protein polymorphisms. *Genetics*, 1974, 77, 343-384.
[21] *Berger E. M., Weber L.* The ribosomes of *Drosophila*. II. Studies on intraspecific variation. *Genetics*, 1974, 78, 1173-1183.
[22] *Brown A. J. L., Langley Ch. H.* Reevaluation of level of genic heterozygosity in natural populations of *Drosophila melanogaster* by two-dimensional electrophoresis. *Proc. Nat. Acad. Sci.* US, 1979, 76, 2381-2384.
[23] *Carson H. L.* The genetics of speciation at the diploid level. *Amer. Natur.*, 1975, 109, 83-92.
[24] *Четвериков С. С. Ж. экспер. Биол.*, 1926, 2-3, I, 26-44.
[25] *Crick F.* Split genes and RNA splicing. *Science*, 1979, 204, 264-271.
[26] *Dobzhansky Th.* Evolution, genetics and man. N. Y., J. Wiley, 1955, 380 p.
[27] *Dobzhansky Th.* Genetics of the evolutionary process. N.Y.-L., Columbia Univ. Press, 1970, 505 p.
[28] *Dobzhansky Th., Ayala F. J., Stebbins G. L., Valentine J. W.* Evolution. San Francisco, W. H. Freeman and Co, 1977, 572 p.
[29] *Дубинин Н. П.* Эволюция популяций и радиация. М., Атомиздат, 1966, 743 с.
[30] *Дубинин Н. П.* Проблема гена в свете молекулярной организации эукариотов и прокариотов. *Успехи соврем. биол.*, 1979, 87, 3, 331-344.
[31] *Дубинин Н. П., Алтухов Ю. П., Сусков И. И.* и др. Экспериментальное обоснование принципов мониторинга генных мутаций у человека. *ДАН СССР*, 1978, 243, 5, 1313-1316.
[32] *Dubinin N. P., Altukhov Yu. P.* Gene mutations (*de novo*) found in electrophoretic studies of blood proteins of infants with anomalous development. *Proc. Nat. Acad. Sci. USA*, 1979, 76, 5226-5229.

[33] *Дуброва Ю. Е.* Влияние стабилизирующего отбора на распределение частот аллелей локусов, кодирующих белки. *ДАН СССР*, 1980, 252, 2, 472-475.
[34] *Harris H.* Enzyme polymorphism in man. *Proc. Roy. Soc. London*, 1966, B, 164, 298-310.
[35] *Harris H., Hopkinson D. A., Robson E. B.* The incidence of rare alleles determining electrophoretic variants : data on 43 enzyme loci in man. *Ann. Hum. Genet.*, 1974, 37, 237-253.
[36] *Hubby J. L., Lewontin R. C.* A molecular approach to the study of genic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*. *Genetics*, 1966, 54, 577-594.
[37] *Jukes T. H.* Comparisons of polypeptide chains of globins. *J. Mol. Evol.*, 1971, 1, 46-62.
[38] *Kimura M., Ohta T.* Mutation and evolution at the molecular level. *Genetics* (suppl.), 1973, 73, 19-35.
[39] *Kirpichnikov V. S.* Genetic basis of fish selection. Springer Verlag, Berlin-Heidelberg-New York, 1981, 410 p.
[40] *Lewontin R. C.* The genetic basis of evolutionary changes. London, 1974, 346 p.
[41] *Lewontin R. C.* Генетическая гетерогенность электрофоретических аллелей. В кн.: XIV Междунар. генет. конгр. Секцион. заседания. Тез. докл. М., Наука, 1978б, I, 465 с.
[42] *Lewontin R. C., Hubby J. L.* A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics*, 1966, 54, 595-609.
[43] *Manwell C., Baker C. M. A.* Molecular biology and the origin of species : heterosis, protein polymorphism and animal breeding. Sidgwick a. Jackson, L., 1970.
[44] *Mayr E.* Animal species and evolution. Oxford Univ. Press, Oxford, 1965.
[45] *Neel J. V.* « Private » genetic variants and the frequency of mutation among South American Indians. *Proc. Nat. Acad. Sci. US*, 1973, 70, 3311-3315.
[46] *Neel J. V., Rothman E. D.* Indirect estimates of mutation rate in tribal American Indians. *Proc. Nat. Acad. Sci. US*, 1978, 75, 5585-5588.
[47] *Nei M.* Genetic distance between populations. *Amer. Natur.*, 1972, 106, 283-292.
[48] *Nei M.* F-statistics and analysis of gene diversity in subdivided populations. *Ann. Hum. Genet.*, 1977, 41, 225-233.
[49] *Nei M., Roychoudhury A. K.* Genic variation within and between the three major races of man, Caucasoids, Negroids and Mongoloids. *Amer. J. Hum. Genet.*, 1974, 26, 421-443.
[50] *Nevo E.* Genetic variation in natural populations : patterns and theory. *Theor. Pop. Biol.*, 1978, 13, 121-177.
[51] *Ohno S.* Evolution by gene duplication. B., Spring.-Verl., 1970, 178 p.
[52] *Омельченко В. Т.* Электрофоретическое исследование гемоглобинов рыб Дальнего Востока. *Генетика*, 1974, 10, 35-43.
[53] *Powell J. R.* The founder-flush speciation theory : an experimental approach. *Evolution*, 1978, 32, 465-474.
[54] *Рычков Ю. Г.* Реакция популяций на изоляцию. В кн.: Проблемы эволюции. Новосибирск, Наука, 1968, I, 212-236.
[55] *Рычков Ю. Г.* Некоторые популяционно-генетические подходы к антропологии Сибири. *Вопр. антропологии*, 1969, 33, 16-33.
[56] *Рычков Ю. Г.* Сравнительное изучение генетического процесса в урбанизированной и изолированной популяциях. *Вопр. антропологии*, 1979, 63, 3-21.

[57] *Рычков Ю. Г., Movsesян А. А.* Генетико-антропологический анализ распределения аномалий черепа у монголоидов Сибири в связи с проблемой их происхождения. *Бюл. МОИП*, 1972, 43, 114-132.

[58] *Tsuyuki H., Ronald A. P.* Existence in salmonid hemoglobins of molecular species with three and four different polypeptides. *J. Fish. Res. Board Canada*, 1970, 27, 1326-1328.

[59] *Witkins N. P.* The subunit composition of the haemoglobins of the Atlantic salmon (*Salmo salar* L.). *Biochim. et biophys. Acta*, 1970, 214, 52-63.

L'ORGANISATION SYSTÉMIQUE DES POPULATIONS ET SON IMPORTANCE POUR LA SAUVEGARDE ET L'UTILISATION RATIONNELLE DU FOND GÉNIQUE DE LA BIOSPHÈRE

Y. Altoukhov et E. Salmenkova

Parmi les différents courants de la science moderne, on voit s'avancer de plus en plus la ligne de recherches réunies dans le cadre du problème « L'Homme et la Biosphère ». Cette tendance est dictée par la vie même, car dans les conditions d'un progrès technique toujours accéléré et de la croissance de la population, l'action exercée sur la Nature prend une ampleur sans précédent dans l'histoire humaine. Dans les différents pays du monde, des efforts sont faits pour stabiliser et améliorer le milieu d'habitation, des mesures sont planifiées pour protéger la biosphère. En 1972, s'est tenue la Session plénière de l'Académie des Sciences (Moscou) à laquelle le problème de la biosphère a été discuté sous tous les aspects. Les principales orientations des recherches ont été indiquées. L'une d'entre elles est directement liée au thème de cet article, il s'agit des questions relatives à l'utilisation rationnelle des ressources biologiques.

A première vue, il peut sembler étrange que, parlant d'un problème écologique si complexe, nous nous ne bornions à l'examen de l'organisation de la vie qu'au niveau des populations. Cependant, cette impression doit se dissiper si l'on se souvient que les *populations sont des groupes d'individus historiquement constitués et s'autoreproduisant* et que c'est la forme d'existence de toutes les espèces biologiques quelles qu'elles soient.

D'après les conceptions actuelles, la *population* est l'unité fondamentale de l'évolution. Mais elle est aussi l'objet principal de notre activité pratique, rappelons-nous la chasse, la sylviculture, la pêche. Actuellement, les populations naturelles ressentent une forte pression anthropogène. Au cours, tout au plus, des cent dernières années, l'activité humaine est parvenue à menacer de disparition 25 000 espèces de plantes supérieures et plus de 1 000 espèces d'animaux vertébrés; plus de 100 races d'animaux domestiques sont au bord de l'extinction (World Conservation Strategy, 1980).

© Издательство ВИНИТИ, 1983.

Quelle en est la raison? Est-ce seulement que la pression de l'exploitation s'est excessivement accrue ou bien les conditions historiquement constituées de la vie se sont-elles violemment modifiées? Peut-être s'agit-il encore de ce que simplement nous n'utilisons pas rationnellement les populations naturelles, que nous perturbons les mécanismes internes naturels qui leur permettent une reproduction stable au fil des générations en maintenant l'équilibre avec le milieu environnant? On peut chercher les réponses à ces questions dans différentes voies de la recherche biologique en utilisant diverses approches et méthodes. Parmi elles, l'approche de la génétique des populations est d'une grande importance.

Il y en a deux raisons. Premièrement, toutes les particularités biologiques des populations sont des dérivées de leurs particularités héréditaires ou, comme disent les généticiens, de leur patrimoine génétique collectif — collection de l'information héréditaire. Deuxièmement, en génétique des populations, il existe depuis longtemps, et elle continue à se perfectionner, une théorie quantitative des facteurs naturels sous la pression desquels l'aspect des populations reste soit inchangé, soit modifié dans les générations successives, ce qui provoque des variations corrélatives des propriétés biologiquement importantes. En d'autres termes, il y a, en génétique des populations, des modèles mathématiques. Ces modèles peuvent être adéquats ou non adéquats à la Nature, mais ils sont intéressants en ce sens qu'ils permettent de planifier d'une certaine façon la recherche. De plus, si nous observons la conformité de la situation naturelle avec le modèle, nous obtenons ainsi la possibilité de faire des prévisions d'une longue portée.

Le but de ce travail est de montrer les faits et les conclusions déduites que l'on obtient par l'étude des populations naturelles à l'aide des méthodes de la génétique. L'attention principale sera accordée aux espèces de poissons d'une valeur économique [1-9, 12, 17-19, 21-24, 26, 27]. En U.R.S.S., ces recherches ont été effectuées depuis 1962 dans le cadre de l'Académie des Sciences (Institut de Biologie marine du Centre scientifique d'Extrême-Orient, Institut de Génétique générale) et à la faculté de biologie de l'Université de Moscou avec le soutien de plusieurs établissements du Ministère de la Pêche de l'U.R.S.S.

Chaque discipline scientifique a son propre objet et sa méthode. Pour que la suite de l'exposé soit claire, il faut caractériser les populations que nous avons choisies — objets des recherches — et exposer l'essence des procédés méthodiques utilisés.

Nous avons étudié les bancs dits locaux: de grandes populations localisées sur des territoires géographiques isolés les uns des autres par des frontières naturelles, barrières constituées au cours de l'histoire des temps et conférant un certain aspect mosaïque à la répartition spatiale de l'espèce dans son aire d'habitation. La particularité

essentielle de l'approche consiste en ce que, en plus des bancs natifs s'autoreproduisant et qui ne sont pas encore détruits entièrement par la pêche —sébaste (*Sebastes mentella* Travin) dans le nord-ouest de l'Atlantique, anchois européen (*Engraulis encrasicholus* L.) dans la mer Noire et la mer d'Azov, saumon rouge du Pacifique (*Oncorhynchus nerka* Walb.) dans les lacs de frai du Kamtchatka — , on a également étudié des bancs artificiellement entretenus dans des usines de pisciculture (saumon kéta de Sakhaline (*Oncorhynchus keta* Walb.)).

En étudiant le sébaste dans le nord-ouest de l'Atlantique, l'anchois dans le bassin de la mer Noire et de la mer d'Azov et les salmonidés du Pacifique en Extrême-Orient, nous avons spécialement choisi des espèces de biologie différente, car il importait d'élucider comment influe cette circonstance sur la reproductivité des populations, quelles sont les particularités de leur organisation génétique interne? Le sébaste est benthique, l'anchois qui vit dans toute l'épaisseur de l'eau est un poisson très mobile, les salmonidés passent une partie de leur vie dans l'océan où ils arrivent à maturité et grossissent, mais vont se reproduire dans les eaux douces. Cette diversité des objets nous permettait d'espérer de mettre en évidence aussi bien certaines particularités que les causalités générales de leur organisation au niveau des populations.

D'une autre part, afin d'étudier la génétique des populations, il faut pouvoir analyser les gènes, et, d'ordinaire, les travaux correspondants se poursuivent dans les cas où l'action du gène se projette sur un caractère extérieur, bien visible, et définit la variabilité individuelle discontinue des exemplaires au sein de l'espèce. Ce phénomène porte le nom de *polymorphisme héréditaire*. Mais chez de nombreuses espèces, il n'y a pas de polymorphisme, elles ont un aspect uniforme, sont monomorphes, leurs gènes sont cachés à l'observateur et, par conséquent, l'étude génétique est impossible.

Comment surmonter un obstacle de ce genre?

Le développement de la biologie moléculaire permet maintenant d'étudier la génétique de n'importe quelle espèce en s'appuyant sur les particularités héréditaires cachées, « inscrites » dans les structures protéiques de l'organisme. Ces particularités sont mises en évidence par des procédés divers: tests immunologiques ou électrophorèse des protéines.

La figure 1 montre plusieurs systèmes de polymorphisme génétique des protéines que nous avons découvert chez les salmonidés du Pacifique par électrophorèse [19, 23, 24, 26, 27]. La méthode consiste en ce qui suit: le sang ou les extraits de différents tissus sont placés dans des appareils spéciaux et dans les milieux d'entretien neutres (gels), on trie les molécules dans le champ électrique selon la charge ou d'autres caractères déterminés par l'hérédité. De cette façon, dans n'importe quel échantillon, dans n'importe quelle popu-

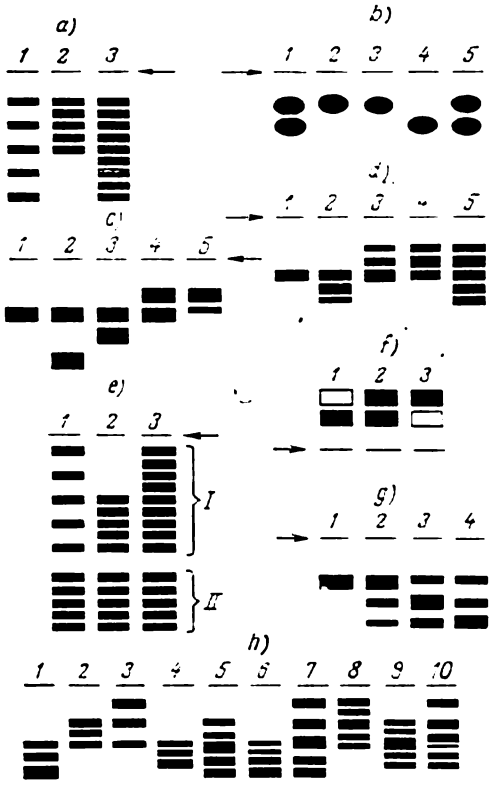


Fig. 1. Variantes électrophorétiques héréditaires de protéines des salmonidés du Pacifique:

a — lactate-déshydrogénase (LDH) sérique du saumon nerka: *1* et *2* — respectivement homozygotes B'B' et BB, *3* — hétérozygote B'B; *b* — phosphoglucomutase (PGM) musculaire du saumon nerka: *2*, *3* et *4* — respectivement homozygotes AA et BB, *1* et *5* — hétérozygotes AB; *c* — phénotypes de l'albumine (Alb) sérique du saumon kéta: *1* — C, *2* — AC, *3* — BC, *4* — CD, *5* — D; *d* — phénotypes de la malate-déshydrogénase (MDH) musculaire du saumon kéta: *1* — phénotype A, *2* — B, *3* — C, *4* — C', *5* — D; *e* — génotypes de la LDH musculaire du saumon kéta: *1* et *2* — respectivement homozygotes AA et A'A'; *3* — hétérozygote A'A; *f* — phénotypes de la LDH sérique du saumon kéta: *1* — F, *2* — FS, *3* — S; *g* — génotypes de l'aspartate-aminotransférase (ASAT) musculaire du saumon kéta: *1* — AAAA, *2* — AAAA', *3* — AAA'A', *4* — AA'A'A'; *h* — génotypes de l'isocitrate-déshydrogénase (IDH) hépatique du saumon kéta: *1* — FF, *2* — SS, *3* — S'S', *4* — F'F', *5* — FS, *6* — FF', *7* — FS', *8* — S'S', *9* — F'S, *10* — F'S'

lation d'une seule et même espèce on peut découvrir une variabilité héréditaire individuelle.

Il suffit de jeter un coup d'œil sur le schéma présenté sur la figure 1, *b*, où l'on voit les différences entre les individus du saumon rouge du Pacifique d'après les gènes allèles gouvernant la synthèse de la phosphoglucomutase. Voici comment se déchiffre ce tableau: un gène est une raie sur l'électrophorégramme, les numéros diffé-

rents indiquent des poissons différents tirés de la même population. Chez les poissons n°s 2-3, on trouve un gène; chez le n° 4, un autre. Les n°s 1 et 5 portent les deux gènes, ce sont des hétérozygotes: types intermédiaires entre deux premiers génotypes qui sont homozygotes. En principe, la même interprétation, ou une autre rapprochée, est juste aussi pour les autres exemples de polymorphisme donnés sur la figure 1. Ainsi, du point de vue de la génétique, une population est un mélange de génotypes. De nombreuses particularités génétiques des populations et, avant tout, l'effectif et la productivité, dépendent de cette structure génétique qui leur est propre. Si le rapport des génotypes et, plus exactement, la fréquence des gènes dans la population sont invariants de génération en génération, la population est en équilibre. Si les rapports varient, on parle d'une population qui évolue. Ce sont les processus dits de microévolution décrits sous leur aspect le plus général dans la théorie de la génétique des populations par les expressions mathématiques présentées sur la figure 2.

Pour ce qui va suivre, il est important d'avoir une idée des *populations panmictique* et *subdivisée*. Dans la partie supérieure de la figure 2, on a représenté schématiquement une population panmictique dans laquelle les individus, qui la composent, se croisent librement entre eux. Si des conditions déterminées sont observées (effectif illimité, absence d'action de facteurs extérieurs, etc.), la population panmictique reste en équilibre, stable au fil des générations. Ceci s'explique par la *loi de Hardy-Weinberg* connue.

Mais des populations idéales de ce genre — d'effectif illimité et existant hors de l'action du milieu — ne sont qu'une abstraction étant donné que dans la Nature agissent sans cesse des facteurs perturbant leur équilibre génotypique. Par exemple, la sélection naturelle peut agir et avoir pour corollaire le fait que certains génotypes laissent davantage de descendants que d'autres et ceci entraîne une modification de la structure génétique de la population. Il peut y avoir *migration* — flux ou reflux — des gènes, et la population changera également sa caractéristique génétique. Enfin, il peut se produire ce qu'on appelle *dérive génétique*: l'effectif de la population baisse fortement et elle « se trompe » lors de la reproduction du patrimoine génétique de la génération suivante.

On peut s'imaginer que tous ces facteurs agissent simultanément sur la population, ce qui doit entraîner une complication de sa structure interne. La population panmictique se divisera alors en un ensemble de *subpopulations* reliées par tel ou tel degré d'interaction migratoire. La distribution de fréquences géniques dans cette population subdivisée redevient stable, fixe en raison des tendances centrifuges (dérive) et centripètes (migration) qui s'équilibrent réciproquement.

S. Wright [50, 51], R. Fisher [35] et J. Haldane [36] ont fait

un grand apport à la justification de cette théorie. En U.R.S.S., mentionnons les travaux fondamentaux dans cette direction de S. Tchetverikov [48], N. Dubinine et D. Romachov [32, 34].

Examinons maintenant à quoi nous avons affaire dans la Nature?

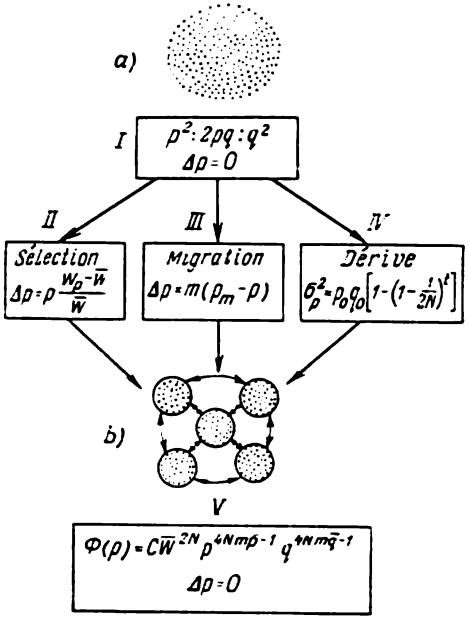


Fig. 2. Principaux modèles mathématiques en génétique des populations:
a — population panmictique d'effectif illimité restant en équilibre ($\Delta p = 0$) en l'absence de sélection (W), de migration (m) et de dérive génétique (N) — loi de Castle-Hardy-Weinberg décrite par la formule *I* où p et q sont des fréquences de gènes associés par hasard dans les génotypes homozygotes (p^2 et q^2) et hétérozygotes (pq); *II*, *III* et *IV* — formules élémentaires correspondant respectivement à la sélection, la migration, la dérive en tant que facteurs principaux déterminant les variations des fréquences géniques (Δp); *b* — population subdivisée formée sous l'action simultanée de la sélection, de la migration et de la dérive des gènes sur la structure génétique; *V* — fonction mathématique caractérisant la distribution de fréquences géniques dans une telle population

La figure 3 montre comment peut être caractérisée toute population naturelle du point de vue de la génétique et de la biologie. En qualité d'exemple, examinons un banc de sébastes (*Sebastes mentella*) localisé près de Terre-Neuve dans le nord-ouest de l'Atlantique et isolé des autres bancs par les particularités de l'hydrographie de l'aire d'habitation. On voit à gauche Terre-Neuve et à droite la chute des profondeurs continentales où s'accumulent les sébastes, une traînée continue de poissons cachée à nos yeux par l'épaisseur de l'eau de mer. Les points marquent la localisation des stations de chalutage (v. fig. 3, *a*). Le bateau de recherche descend de haut en

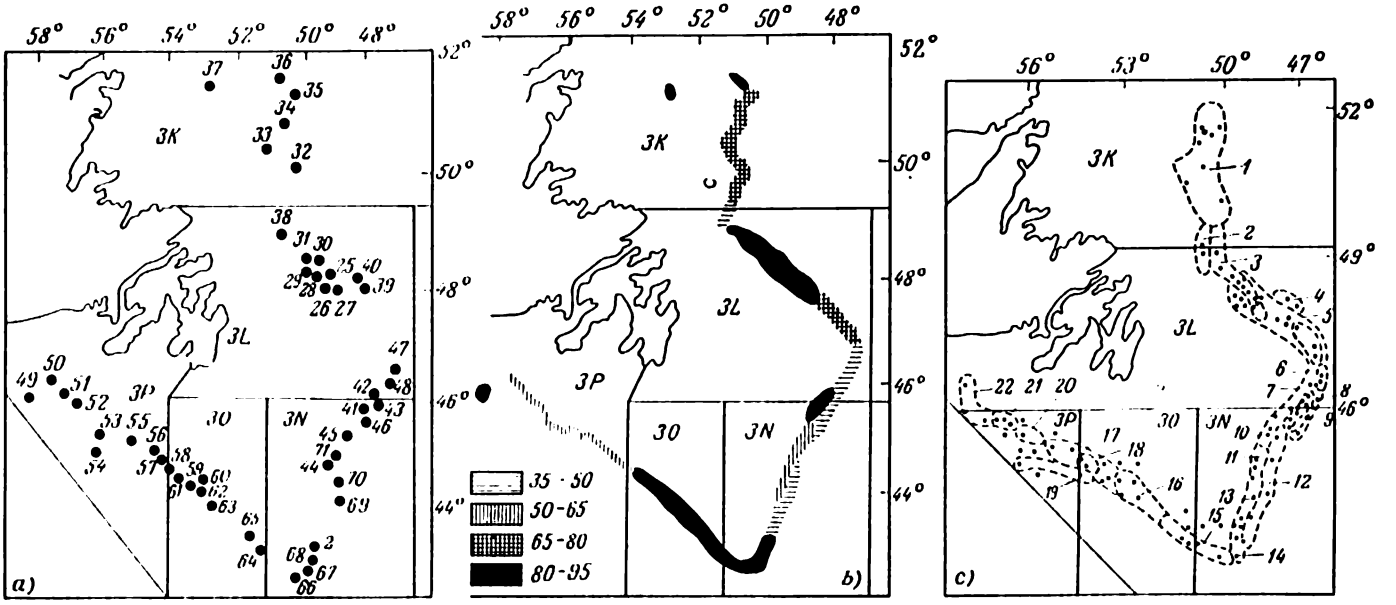


Fig. 3. Hétérogénéité héréditaire d'un banc de sébastes (*Sebastes mentella*) localisé dans la région de Terre-Neuve: *a* — schéma de localisation des stations de chalutage; *b* — carte génogéographique montrant la variabilité spatiale de la fréquence du gène des groupes sanguins; c — localisation spatiale de subpopulations différant génétiquement et constituant la structure reproductrice du banc. Les contours délimitent l'ensemble des échantillons qui, différant les uns des autres, se caractérisent par un maximum d'homogénéité biologique

bas, il drague sans cesse, et de chaque filet de dragage le poisson est pris pour le classer selon ses caractères héréditaires (dans le cas envisagé, il s'agit des groupes sanguins) et ses caractères biologiques, tels que la longueur du corps, le sexe, l'âge, le stade de maturité sexuelle.

On découvre une hétérogénéité héréditaire fortement prononcée du banc, ce que représente la *carte génogéographique* (v. fig. 3, *b*). Nous voyons les champs de différentes fréquences géniques ou concentrations, témoignage indirect de ce que le banc étudié peut être différencié en de plus petits bancs génétiquement différents. En effet, si l'on analyse d'une façon déterminée la variabilité des caractères biologiques mentionnés, on peut découvrir dans les limites du banc un ensemble de groupements biologiquement homogènes, mais, en même temps, différant entre eux. Ce sont les *populations élémentaires* ou *subpopulations* dont chacune est caractérisée par un patrimoine génétique spécifique (v. fig. 3, *c*).

Le schéma-carte de la figure 4 représente la structure de populations d'un banc local de saumons rouges du Pacifique, qui se reproduisent dans le lac Azabatchi (Kamtchatka). Les points désignent les divers endroits de frai occupés par différentes populations élémentaires. Cette structure peut être observée directement, aussi n'est-il pas nécessaire de se livrer à de laborieuses figurations des populations.

Quel est le sens de cette différenciation de l'espèce ou d'un de ses lots, le banc local? Est-ce que ce sont des populations sans rapport entre elles ou les parties d'un tout unique qu'est le banc? Les observations écologiques montrent que de telles subpopulations restent, vraiment, parties du tout, étant donné que pas une d'elles n'englobe toute la diversité héréditaire propre au banc et, en outre, elles interagissent constamment entre elles par échange de gènes. Donc, un banc local pris à part nous apparaît comme un certain système de subpopulations liées par une interaction.

Pour une étude plus approfondie des phénomènes génétiques dans une telle population, il est important de la comparer au modèle mathématique le plus adéquat. Souvenons-nous qu'il s'agissait plus haut de deux types de modèles: la population panmictique et la population subdivisée. Il est clair que nous devons ici choisir le modèle de la population subdivisée et faire les calculs mathématiques. En étudiant en détail pendant plusieurs années un banc de saumons rouges au Kamtchatka d'après les enzymes génétiquement polymorphes, nous avons découvert une correspondance satisfaisante entre la situation naturelle et le modèle « insulaire » de structure de populations proposé par S. Wright.

La figure 5 montre l'éventail de la variabilité génétique réelle des populations (histogrammes) comparée aux distributions théoriques des fréquences géniques de la lactate-déshydrogénase et de la

phosphoglucomutase (courbes). C'est la partie inférieure de la figure (*b*) qui offre le plus d'intérêt. Elle montre la distribution des fréquences géniques de la PGM dans 120 subpopulations constituant la structure reproductrice du banc. La première courbe ne correspond

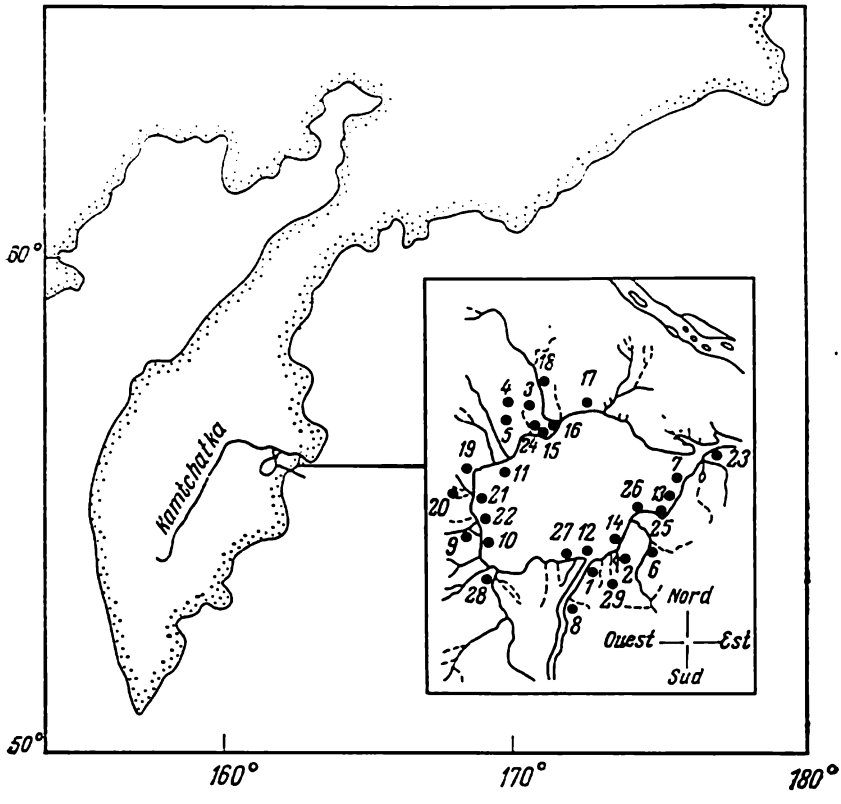


Fig. 4. Localisation des subpopulations de frai (nos *1-29*) constituant la structure reproductrice d'une population isolée de saumon rouge (*Oncorhynchus nerka* Walb.) du lac Azabatchi (Kamtchatka)

strictement pas à la distribution empirique. Dans ce cas, comme dans l'analyse de la distribution des fréquences géniques de la LDH (v. fig. 5, *a*), on suppose que l'état stationnaire du système est assuré par l'équilibrage des pressions de migration et de dérive génétique due au faible effectif de certaines subpopulations. Cependant, une telle supposition ne suffit pas. De plus, la seconde courbe correspond bien aux données empiriques. Ici, le facteur sélection est introduit dans le modèle et il est orienté en faveur des génotypes hétérozygotes [5, 23, 24].

Malgré la différenciation des populations élémentaires, malgré leur variabilité génétique dans le temps comme dans l'espace, le banc dans l'ensemble reste stable (*Tableau 1*) et, ce qui est particulièrement important, cette stabilité reflète son adaptation au milieu concret, tout au moins en ce qui concerne les génotypes du locus PGM.

En s'appuyant sur ces données, on peut en effet tirer une conclusion que le banc local étudié (comme les bancs d'autres espèces

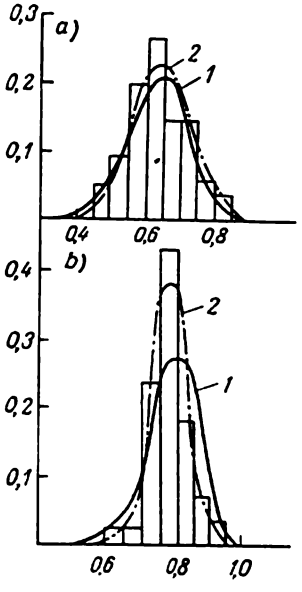


Fig. 5. Distribution des subpopulations de saumon rouge dans la gamme de fréquences géniques de la lactate-déshydrogénase (*a*) et de la phosphoglucomutase (*b*) :

Les histogrammes sont les distributions empiriques ; les courbes, les distributions théoriques :

$$1 - \Phi(q) = C q^{4Nm\bar{q}-1}(1-q)^{4Nm\bar{p}-1};$$

$$2 - \Phi(q) = C(\overline{W})^{2N} q^{4Nm\bar{q}-1}(1-q)^{4Nm\bar{p}-1}.$$

Sur l'axe des abscisses, fréquence génique ; sur l'axe des ordonnées, part de la population (en % du nombre total d'individus étudiés) entrant dans la gamme correspondante de fréquences.

Afin de construire les courbes théoriques on a utilisé les valeurs : $N = 200$ et $m = 0{,}02$. Valeurs du coefficient d'adaptabilité des génotypes pour le locus PGM (W) : $W_{AA} = 0{,}94$; $W_{AB} = 1{,}00$; $W_{BB} = 0{,}76$

de salmonidés) représente un système de populations qui s'est constitué historiquement et stabilisé génétiquement et qui peut rester en équilibre jusqu'à ce que sa structure interne soit soumise à des facteurs extérieurs excessifs.

Cette conclusion se trouve juste non seulement pour les salmonidés, mais aussi pour d'autres poissons et, apparemment, pour n'importe quelles espèces se reproduisant par croisement [10]. Si l'on applique l'approche comparative correspondante, on découvre une stabilité de longue durée des populations subdivisées dans le temps malgré la variabilité affectant leurs parties. Ce fait peut être illustré par plusieurs exemples.

La figure 6, *a* montre la structure génétique de trois générations successives de l'anchois de la mer d'Azov, pris comme un tout. Sur l'axe vertical, on a porté la fréquence génique en tant que caractéristique essentielle en génétique des populations ; sur l'axe horizontal, les échantillons tirés de quelques subpopulations constituant la

structure du banc dans la première, la seconde et la troisième générations successives. Les traits parallèles à l'axe inférieur représentent la fréquence génique dans les échantillons. Les colonnes verticales sont les erreurs types doublées caractérisant la fiabilité statistique

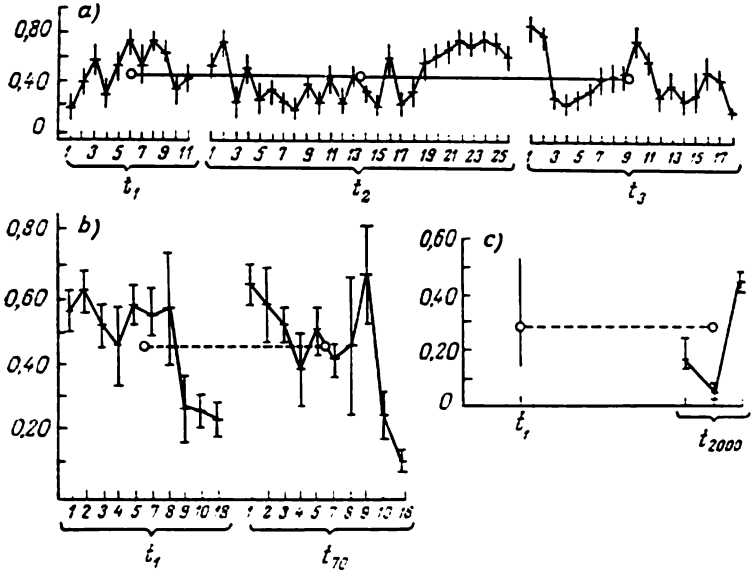


Fig. 6. Exemples de stabilité génétique des systèmes de populations dans le temps:

a — stabilité génétique d'une population d'anchois de la mer d'Azov sur un fond de la variabilité de différentes composantes structurales dans les générations successives t_1, t_2, t_3; *b* — stabilité génétique d'un système de populations de drosophile au fil de 70 générations sur un fond de la variabilité de différentes composantes structurales (d'après Dobzhansky); *c* — stabilité génétique d'un système de populations de littorine au fil de 2000 générations; t_1 — population ancestrale (*grosso modo*: $pA = 0{,}287 \pm 0{,}029$; $N = 479$ individus); t_{2000} — population actuelle représentée par 3 subpopulations ($pA = 0{,}273 \pm 0{,}01$; $N = 3818$ individus).

Sur l'axe vertical, fréquence génique en tant que paramètre principal en génétique des populations; sur l'axe horizontal, échantillons tirés de plusieurs populations élémentaires constituant la structure des systèmes dans les générations successives. Les traits horizontaux représentent la fréquence génique dans les échantillons; les traits verticaux, les erreurs types doublées; les ronds, les moyennes de la fréquence génique pour le système en tant que tout

des données. Nous voyons que malgré la variabilité dans les subpopulations, leur ensemble se maintient avec stabilité comme un tout dans la suite de générations étudiée.

Le deuxième exemple (v. fig. 6, *b*) se rapporte à une autre espèce, la drosophile, dont la génétique des populations a été largement étudiée grâce aux travaux de l'école de Th. Dobzhansky [31]. En appliquant la même approche au matériel correspondant, on peut

Tableau 1

Etat stationnaire des paramètres principaux de la génétique des populations pour un banc de saumons rouges (*Oncorhynchus nerka* Walb.) du lac Azabatchi (Kamtchatka)

Locus	Année	k	n	$\bar{q}$	σ_q^2	Test à l'homogénéité des dispersions (critère de Bartlett)	Test à l'homogénéité du banc (avec degrés de liberté)
LDH	1971	14	737	0,65	0,0074	$2,70 < \chi^2_{0,05}(6) = 12,59$	$50,18 > \chi^2_{0,001}(13) = 34,50$
	1972	20	1022	0,61	0,0048		$37,53 > \chi^2_{0,001}(19) = 36,19$
	1973	21	1007	0,65	0,0078		$65,11 > \chi^2_{0,001}(20) = 45,30$
	1974	14	676	0,66	0,0076		$45,59 > \chi^2_{0,001}(13) = 34,50$
	1977	19	934	0,65	0,0084		$69,34 > \chi^2_{0,001}(18) = 42,30$
	1978	23	1244	0,65	0,0071		$76,62 > \chi^2_{0,001}(22) = 48,30$
	1979	14	846	0,65	0,0091		$67,30 > \chi^2_{0,001}(13) = 34,50$
	Total	125	6466	0,65	0,0076		$427,91 > \chi^2_{0,001}(124) = 160$
PGM	1971	14	748	0,79	0,0028	$9,21 < \chi^2_{0,05}(6) = 12,59$	$22,57 > \chi^2_{0,05}(13) = 22,35$
	1972	20	1004	0,77	0,0019		$21,76 < \chi^2_{0,25}(19) = 22,72$
	1973	21	981	0,79	0,0041		$41,36 > \chi^2_{0,01}(20) = 37,57$
	1974	12	546	0,78	0,0037		$23,02 > \chi^2_{0,01}(11) = 19,68$
	1977	19	936	0,81	0,0032		$39,21 > \chi^2_{0,01}(18) = 34,81$
	1978	23	1241	0,77	0,0025		$35,78 > \chi^2_{0,05}(22) = 33,92$
	1979	14	846	0,78	0,0018		$17,05 < \chi^2_{0,05}(13) = 22,36$
	Total	123	6302	0,78	0,0033		$247,96 > \chi^2_{0,001}(123) = 150,50$

Remarque. k est le nombre de subpopulations,
n nombre d'individus étudiés,
$\bar{q}$ fréquence génique moyenne,
σ_q^2 variance des fréquences géniques entre les populations.

montrer également la stabilité de cette population, considérée ordinairement comme un objet d'une transformation évolutive. Ici, la stabilité génétique est suivie sur 70 générations.

Enfin la dernière illustration : nos données concernant un mollusque, la littorine, habitant la lagune de Boussé dans le sud de Sakhaline (fig. 6, *c*). Cet objet a ceci d'intéressant que, chez lui, l'action du gène est « inscrite » à la surface de la coquille sous forme d'un dessin déterminé qui se maintient chez les fossiles. Son âge est d'environ 5000-6000 ans et, par conséquent, compte tenu de l'âge moyen de reproduction de la littorine, la population actuelle est séparée de l'ancestrale par un intervalle de l'ordre de deux mille générations successives. Nous voyons que, bien que vivant dans les conditions modifiées d'une lagune très isolée, la population de l'espèce étudiée a conservé toute son information héréditaire. On parvient à simuler le même phénomène sur des populations subdivisées expérimentales de drosophile [12-16, 25].

Ainsi, l'étude de la génétique et de la biologie des populations naturelles met en évidence leur organisation systémique, ce qui a une importance non seulement théorique, mais aussi pratique. On comprend bien que si nous avons affaire à un système, nous devons l'aborder comme un tout compte tenu de sa diversité interne. Or, il est facile de montrer qu'ordinairement, la pêche ne tient pas compte de l'organisation interne des bancs, ce qui entraîne la destruction par parties de ces systèmes de populations. Pour illustrer cette thèse, adressons-nous au tableau typique de la pêche en mer dont la stratégie, comme on sait, comporte deux actions essentielles : la reconnaissance des agglomérations assez puissantes de poissons par un bateau spécial et, une fois qu'on les a découvertes, leur pêche par une flottille de chalutiers. Afin de montrer ce tableau dans sa dynamique, on peut le représenter sous forme d'images successives comme on le fait sur la figure 7 correspondant à la figure 3, *c* sur laquelle était démontrée la chaîne mise en évidence par nous de subpopulations de sébastes (*Sebastes mentella*) différant génétiquement. On comprend qu'avec ce genre de pêche, quand les bateaux recherchent chaque fois les secteurs de l'aire d'habitation où l'accumulation de poissons est caractérisée par une densité maximale, certaines subpopulations sont pêchées en excès, d'autres insuffisamment, et, finalement, il se produit une perturbation des canaux formés naturellement de liaison migratoire entre les éléments du système. On n'aurait pu l'éviter qu'en attrapant régulièrement le banc tout entier, compte tenu obligatoirement de sa structure de populations spatiale.

Les mêmes effets néfastes, rattachés à la méconnaissance de la structure de l'espèce au niveau de la génétique des populations, sont observés non seulement dans la pêche. On peut les observer même dans le cas où sont poursuivis des buts visant à la reproduction arti-

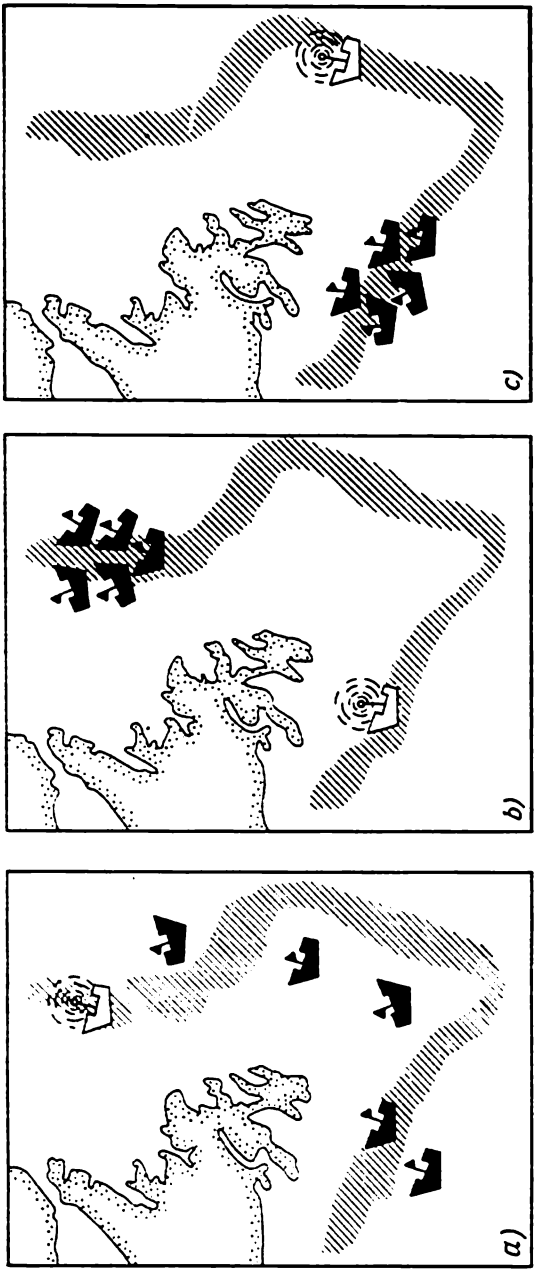


Fig. 7. Schéma approché de la pêche en océan, les flottilles de chalutiers se déplaçant continuellement dans les régions d'agglomération de sébastes découvertes par le bateau de reconnaissance (v. explications dans le texte)

ficielle des ressources biologiques, comme cela a lieu, par exemple, en pisciculture.

La construction d'hydrocentrales, la destruction du milieu environnant et d'autres facteurs sont responsables de ce fait que les populations de certaines espèces de poissons remontant les rivières pour le frai se sont vues privées de leurs sites de frai naturels et qu'on les reproduit artificiellement en mettant en incubation des œufs fécondés dans les bacs des usines de pisciculture, en faisant grossir les juvéniles et en rejetant les fretins dans la mer. La pisciculture des salmonidés a déjà pris, par exemple, une ampleur industrielle. L'année dernière, seulement dans les usines d'Extrême-Orient, on a mis en incubation plus de 800 millions d'œufs de saumon kéta et de saumon gorbuscha. Cependant, malgré le volume croissant des entreprises piscicoles, leur efficacité, dans certains cas, n'est pas élevée.

On est arrivé à élucider [3, 26] qu'ici également, la chose est en rapport avec la négligence de la structure des espèces au niveau de la génétique des populations. D'une part, cela était dû à la pêche intensive des salmonidés dans l'océan et, d'autre part, aux mesures d'acclimatation prises, aux transferts d'œufs artificiellement fécondés d'un fleuve dans un autre. Ces mesures ont été pratiquées à cause de la réduction de l'effectif de « propres » bancs, ce qui a amené certaines usines de pisciculture à utiliser des bancs « étrangers ». L'intensité de telles mesures était assez élevée dans le cas du saumon kéta (*Oncorhynchus keta*), et seulement en quelques années, du fleuve Kalininka (littoral occidental de Sakhaline) dans le fleuve Naïba (littoral oriental de Sakhaline) plus de 360 millions d'œufs ont été transportés, alors que 208 millions d'œufs de saumon kéta de Naïba étaient transférés par avion à l'usine de pisciculture Téplovskaïa (bassin de l'Amour) [26] (*Tableau 2*).

Etant donné qu'on a trouvé des différences génétiques entre ces populations selon l'ensemble des marqueurs biologiques [26] permettant de différencier les populations également dans les agglomérations mixtes, nous avons pu suivre le sort des mesures d'acclimatation (*Tableau 3*). Elles se sont avérées pratiquement inopérantes. Le transfert de centaines de millions d'œufs de poissons de Sakhaline dans l'Amour n'a donné aucun résultat. A Sakhaline, si l'on a pu observer une récupération, c'est seulement dans la première génération : les coefficients de récupération * — indice intégral de l'adaptabilité de la population — se sont avérés de beaucoup inférieurs à ceux observés dans la rivière natale et, par la suite, toute trace de poissons transplantés manquait totalement. Bien plus, la récupération des

* Rapport entre la quantité de reproducteurs revenus dans le fleuve et la quantité de fretins rejetés par l'usine de pisciculture.

Tableau 2

Transports d'acclimatation des œufs de saumon kéta dans la région de Sakhaline, des Kouriles et de l'Amour [26]

Année	Quantité d'œufs transportés, en millions	Année	Quantité d'œufs transportés, en millions
Fleuves		**Fleuves**	
Naïba *-Kalininka		Kalininka-Naïba	
1956	4,500	1959	4,830
1960	2,638	1964	70,610
Naïba-Amour		1965	70,000
1964	22,000	1966	54,123
1967	9,000	1967	25,665
1968	98,000	1968	26,310
1969	62,000	1969	35,563
1970	17,600	1970	18,820
Naïba-Tym'		1971	56,000
1967	9,487	Kourilka-Naïba	
Naïba-Bouyouklinka		1972	30,000
1968	5,390	1973	3,500
1976	3,000	Fleuves de l'île Itouroup-Kalininka	
		1976	environ 30,000
		Tym'-Naïba	
		1966	10,208

* Le fleuve donneur est désigné en premier ; le fleuve récepteur en second.

populations des fleuves Kalininka et Naïba diminuait aussi dans les générations auxquelles on avait enlevé une importante part du patrimoine génétique en vue de les acclimater à un autre secteur de l'aire d'habitation. Ceci se fit sentir surtout sur la population du fleuve Naïba, l'effectif de laquelle diminua fortement à partir de 1971 [3, 8].

Du point de vue de la *génétique des populations, l'acclimatation est une adaptation à un milieu nouveau.* On ne peut juger de son succès qu'après que s'est formée une population s'autoreproduisant avec un patrimoine génétique stable et intégré et qui est capable de subsister un temps indéfini au fil des générations. Malheureusement, on a, aujourd'hui, trop peu de preuves de phénomènes de ce genre.

Pourtant, il y a bien des exemples montrant le contraire. Ainsi, W. Ricker [43] a généralisé un très grand matériel sur les transferts en vue d'une acclimatation effectués surtout sur les populations nord-américaines de salmonidés et a montré que la récupération de la première génération en fleuve étranger avait souvent lieu, mais dans

Tableau 3

Fréquence des gènes de la lactate-déshydrogénase (LDH-A_2), de l'albumine (Alb), de l'isocitrate-déshydrogénase (s-IDH), de l'aspartate-aminotransférase (ASAT), ainsi que des phénotypes de la malate-déshydrogénase (MDH) dans les populations de frai du saumon kéta

Populations observées (fleuves)	LDH-A_2		Alb		s-IDH					ASAT		MDH	
	n	*pA*	*n*	*pD*	*n*	*pF*	*qS*	*rS'*	*rSF'*	*n*	*pA*	*n*	*pA*
Kalininka	1271	0,990	1328	0,852	229	0,667	0,264	0,067	0,002	424	0,920	1641	0,668
Naïba	2185	0,946	2229	0,624	441	0,415	0,430	0,100	0,055	498	0,955	2493	0,875
Tym'	514	0,946	380	0,763						54	0,954	721	0,989
Poronaï	88	0,955			86	0,517	0,343	0,046	0,064	88	0,926	88	0,977
Kourilka	412	0,836	386	0,663	124	0,548	0,292	0,160	0,000	124	0,956	398	0,942
Reïdovaïa	193	0,749	151	0,414								196	0,923
Amour (lac Téploïe)	412	0,993										434	0,998
Amour (liman)	412	0,998										412	0,995

une mesure beaucoup moindre qu'en fleuve natal. La plupart du temps, la récupération dans les générations suivantes baisse fortement ou cesse tout à fait. Le saumon gorbuscha extrême-oriental acclimaté en fleuves de la presqu'île de Kola depuis 1956 par des transferts presque annuels d'œufs de saumon gorbuscha de l'île de Sakhaline fournit de faibles récupérations, généralement dans les lignées des années impaires seulement [30]. Pourtant, il est évident que cette nouvelle population ne se reproduit presque pas naturellement.

Tous ces faits et de nombreux autres parlent de l'unicité et du conservatisme des adaptations locales formées par la sélection au cours de milliers de générations dans un milieu concret auquel se rattache toute l'histoire naturelle de telle ou telle population.

Quand nous perturbons les liaisons qui se sont constituées au cours des siècles, que nous arrachons une population à son propre milieu historiquement formé et que nous la transportons dans un autre milieu, la réserve de solidité génétique ne suffit généralement pas.

Cependant, il est nécessaire de tenir compte de ce qu'une acclimatation réussie s'avère, en règle générale, indésirable à cause des séquelles écologiques: il y a alors une élimination des populations des espèces locales, ayant pour base une concurrence quant à la nourriture et à la reproduction, ou bien, ce qui, apparemment, a lieu le plus souvent, une propagation de vecteurs de maladies dangereuses contre lesquelles les espèces locales n'ont pas élaboré d'immunité. Pour illustrer ce que nous venons de dire, mentionnons une tentative oubliée, faite il y a plusieurs dizaines d'années, pour installer dans la mer d'Aral l'esturgeon étoilé de la Caspienne. Cette mesure trouvait sa justification dans le fait que, dans la mer d'Aral, à ce qu'il semblait, étaient réunies toutes les conditions favorisant l'augmentation de l'effectif des esturgeons, car depuis toujours une de leurs espèces, l'esturgeon bâtard, y habitait. Cependant, cette tentative se solda par une défaite totale: non seulement l'esturgeon étoilé ne s'adapta pas aux circonstances nouvelles, mais il contribua à la perte de la population de l'esturgeon bâtard en l'infectant par un protozoaire parasitaire, la nitzschia. Pour l'esturgeon étoilé, ce parasite n'était pas dangereux, car au cours d'une longue cohabitation évolutive, l'esturgeon avait pu élaborer une immunité. Or, au cours de son histoire, l'esturgeon bâtard n'avait jamais rencontré la nitzschia.

On pourrait augmenter à l'infini le nombre de ces exemples [12, 47, 49], mais nous n'en parlons que très rarement. Cependant, beaucoup de mesures d'acclimatation, qui n'ont pas encore subi de vérification dans la vie, deviennent souvent l'apanage de la grande presse et créent une fausse impression de succès là où, en réalité, il n'y en a pas.

Il faut reconnaître que, malgré les résultats scientifiques appré-

ciables dans le domaine de l'acclimatation et l'attitude nette de l'Union internationale de la protection de la nature (World Conservation Strategy, 1980, v. également [29]) qui considère l'acclimatation comme le principal facteur de la pollution biologique de l'environnement, il n'y a pas encore en biologie à ce sujet l'unanimité qu'il faudrait. La destruction de fonds géniques de la biosphère, uniques et d'une haute valeur, se poursuit.

Est-ce que cela signifierait qu'il faut d'une façon générale renoncer à une intrusion quelle qu'elle soit dans la Nature lors de la décision des tâches scientifiques et pratiques qui se dressent devant l'homme? Non, apparemment, notre pratique agricole en est un exemple: sans agriculture développée l'humanité actuelle ne pourrait exister. Il y a, cependant, une différence essentielle entre les conditions de vie des populations agricoles et des populations naturelles. Dans la Nature, les espèces vivent dans un milieu varié, hétérogène, non soumis au pouvoir de l'homme, alors que les races d'animaux domestiques et les variétés de plantes existent dans des conditions contrôlées par l'homme dans une grande mesure. La stabilité du milieu multipliée par les qualités héréditaires productives des populations agricoles permet de réaliser ce type d'exploitation dirigé et prévu que l'on peut appeler intensif.

Compte tenu des succès de la génétique des populations naturelles, nous pouvons implanter dans notre activité pratique des procédés de leur utilisation industrielle, de reproduction artificielle et d'acclimatation tels que nous pouvons passer à un type d'exploitation intensif. Mais cela n'est réalisable que dans le cas où nous tiendrons compte des particularités de l'organisation systémique des populations et où nous maintiendrons leur hétérogénéité héréditaire formée au cours de l'histoire, que si se conservent les mécanismes d'autorégulation qui assurent leur adaptation efficace aux conditions du milieu naturel variant normalement. *C'est seulement dans cette voie qu'est possible la réalisation pratique d'une stratégie optimale d'exploitation ayant pour but non seulement de tirer un profit économique, mais aussi de conserver pour une durée illimitée les ressources biologiques.*

Pour organiser une pêche dirigée, il faut avoir une notion nette des particularités de la structure de populations de l'espèce qui nous intéresse, comprendre à quel point sont isolées les subpopulations, connaître les rythmes de leur renouvellement et la dynamique de leur effectif et, enfin, compte tenu de la structure, donner des recommandations argumentées telles que l'intensité de la pêche retirant les poissons de la subpopulation ne surpasse pas les cadences du renouvellement naturel ou artificiel de cette dernière.

La structure de subpopulations doit également être prise en ligne de compte lors de la reproduction artificielle des populations. Il ne faut pas oublier que chacune d'elles a son optimum biologique déterminé par l'évolution et par la situation actuelle de la population

dans l'écosystème. En ce qui concerne l'effectif, les limites de l'optimum sont fixées par ses niveaux minimal et maximal dont les appréciations stables ne peuvent être obtenues qu'en faisant la moyenne sur les résultats d'observations systématiques durant de longues années. On sait bien que sans connaître cette zone de stabilité du système, il est très difficile de le reproduire artificiellement, c'est-à-dire de le diriger.

L'auteur comprend, il va de soi, que les notions exposées contredisent les vues selon lesquelles il est admis d'opposer certains génotypes qui sont « plus intéressants » pour l'homme (un engraissement rapide, une fertilité élevée, les rythmes de maturation, par exemple) à d'autres ayant « moins de valeur » en insistant donc sur l'utilité d'un travail de sélection dirigée des populations naturelles [41, 42]. On sait que sur ce principe repose toute notre pratique agricole. Pourtant, il ne faut pas oublier que le renforcement excessif, au cours de la sélection dirigée, de certains caractères et propriétés est inéluctablement rattaché, en vigueur de l'existence de corrélations négatives dans le système de l'ontogénèse [44], à l'affaiblissement et à la détérioration d'autres et entraîne une diminution de la diversité génétique de la race ou de la variété, et que c'est le « prix » ordinaire de la sélection.

De regrettables exemples de ce genre sont nombreux même dans la production agricole avec ses possibilités particulières d'entretenir la stabilité du milieu, souvenons-nous ne serait-ce que de l'échec de la « révolution verte ». En ce qui concerne la pisciculture, les choses ne vont pas mieux. Néanmoins, une expérience négative n'est généralement pas discutée et on en appelle d'ordinaire aux mêmes exemples plus ou moins réussis et qui ne sont pas nombreux. De plus, ce sont des effets passagers à l'échelle des intervalles de temps qui caractérisent la durée de vie des populations natives historiquement constituées. Du point de vue de la génétique, ce ne sont pas les années, mais les générations qui sont la mesure du temps dans la nature vivante. Malgré la critique subie par notre approche [41, 42, v. aussi 7], son importante signification pour l'utilisation rationnelle des ressources biologiques est reconnue par beaucoup de chercheurs [28, 33, 37-40, 46]. De plus, son procédé pratique correspondant est reconnu comme invention [20] et est implanté dans les usines de pisciculture de Sakhaline. D'après les données dont nous disposons, on peut constater son effet positif évident dans plusieurs cas.

Dans les années à venir, il est nécessaire d'apprécier minutieusement l'efficacité économique de l'implantation de recommandations génétiques et de supprimer certaines difficultés objectives faisant obstacle à leur utilisation plus efficace dans la pratique. La difficulté principale consiste en ce que, jusqu'à présent, l'efficacité du travail des usines de pisciculture est évaluée non pas par les coefficients de récupération par la pêche comme mesure de l'adaptabilité

biologique (génétique) de la population, mais par les volumes d'œufs mis en incubation et des juvéniles produits.

Or, à la lumière de ce que nous venons de dire, cette approche ne peut être efficace qu'à condition d'obtenir le point d'équilibre entre l'effectif de la population et les ressources du milieu, avant tout en assurant les alevins s'engraissant en nourriture dans les secteurs marins à proximité des estuaires. Si le point de l'optimum est dépassé, la « résistance du milieu » augmente, ce qui entraîne inévitablement une baisse de l'adaptabilité de la population et la réduction de son effectif. Malheureusement, des recherches de ce genre en rapport avec la rationalisation de la culture des salmonidés manquent jusqu'à présent, bien que leur extrême importance ne soulève pas de doutes. Il importe également de modifier les critères admis de l'évaluation de l'efficacité de la pisciculture dans les bassins naturels et de réunir directement la pêche des bancs à leur reproduction artificielle dans le cadre d'un système unique et consciemment dirigé [20].

Les principes de la génétique des populations sont les mêmes pour toutes les espèces biologiques. Nous avons apporté ici autant d'attention aux populations de poissons seulement parce que, en raison de leur importance pratique, elles sont devenues, au cours des 15 à 20 dernières années, un objet de recherches intensives ayant pour base les positions et les méthodes de la génétique biochimique. A leur tour, ces travaux ont permis de mettre en évidence l'organisation systémique des populations isolées d'autres espèces ayant une écologie notoirement différente, d'obtenir de nouvelles données sur les mécanismes de leur stabilité génétique et d'élucider les causes de la perturbation de cette stabilité.

* * *

Le développement historique a fait en sorte que la génétique des populations, née au début de notre siècle, a rattaché ses intérêts principaux aux problèmes évolutifs de la biologie, concentrant l'essentiel de son attention sur les différents aspects de la dynamique des populations. Ceci a contribué à la création d'une théorie probabiliste minutieusement élaborée de l'évolution, au progrès rapide de nos connaissances et à leur utilisation pratique efficace dont la direction principale était coordonnée à la foi en des facultés transformatrices illimitées de l'homme. Cependant, nous prenons nettement conscience aujourd'hui que dans de nombreux cas les actions sur la biosphère entraînent inéluctablement des situations qu'on ne peut qualifier autrement que comme situations de crise.

Donc, il est impossible de mettre le signe égal entre la mutabilité naturelle inhérente à la vie et les modifications que lui apporte l'activité transformatrice de l'homme.

Il est nécessaire de distinguer ces deux facteurs et de comprendre que les bornes de la stabilité génétique des populations, des espèces et des écosystèmes de même que leurs possibilités d'évolution ne sont pas illimitées.

Les phénomènes funestes rattachés à la diminution de la diversité génétique de la biosphère confirment, comme jamais auparavant, la justesse du sage avertissement d'Engels nous disant qu'à chaque pas les faits nous rappellent que nous n'avons aucunement sur la Nature le pouvoir du conquérant sur un peuple soumis, que nous ne la vainquons pas comme quelqu'un qui serait hors de la Nature, qu'au contraire, nous lui appartenons par notre chair, notre sang, notre cerveau et que nous nous trouvons dans son sein, que toute notre domination sur elle consiste en ce que, à la différence de tous les autres êtres, nous pouvons connaître ses lois et les appliquer correctement.

Les lois de la Nature soulignent la nécessité impérieuse de cesser de polluer l'environnement, d'abattre les forêts, d'épuiser le sol ou de tuer les animaux sauvages dans une mesure dépassant leur faculté naturelle de reproduction. La diversité génétique de la vie, les fonds géniques encore intacts de la biosphère doivent être conservés.

Mais comment y parvenir?

Actuellement, le problème de la protection des fonds géniques se résoud principalement par l'organisation de parcs nationaux et de domaines interdits, c'est-à-dire par la création des réserves de la vie peu modifiées ou, dans l'idéal, nullement modifiées dans les conditions d'un environnement changeant brutalement [45]. Une telle approche est d'une extrême importance. Cependant, il faut également prendre sous un contrôle sévère les canaux de la pression anthropogène sur la biosphère en rapport avec l'exploitation et la reproduction artificielle des ressources biologiques. Dans les phrases précédentes, nous nous sommes efforcés de montrer qu'il n'y a là aussi qu'une seule issue : en s'appuyant sur les mécanismes de la stabilité génétique des populations qui se sont constitués au cours de l'évolution, remanier notre stratégie d'interaction avec la Nature de telle façon qu'elle ne s'accompagne pas de la destruction des fonds géniques de la biosphère et qu'elle contribue non pas à la croissance extensive, mais au développement stable et à la coexistence du système « L'Homme et la Biosphère » dans une suite de générations illimitée. Un tel modèle peut être appelé *modèle de l'optimum socio-écologique.*

Sa réalisation pratique impose un énorme travail coordonné aux représentants de différentes sciences, et la génétique des populations peut et doit y apporter sa contribution. De ce point de vue, ce qui est urgent, c'est d'adopter une approche de la protection et de l'utilisation rationnelle des ressources naturelles qui se fonderait non seulement sur les principes de l'évolution, mais aussi sur les facteurs

et les conditions de la stabilité génétique des populations et des espèces.

La protection de la variété génétique des systèmes de populations encore intacts, la restauration de ceux d'entre eux dont la structure est déjà perturbée et, enfin, la création de nouveaux systèmes de populations, tels sont les problèmes essentiels dont la réalisation est capable d'éliminer ou, tout au moins, d'affaiblir le conflit actuel entre l'humanité et la biosphère.

Mais le même principe de maintien de la diversité génétique constituée au cours de l'évolution doit être juste pour n'importe quels niveaux d'intégration biologique au sens de leur stabilité vis-à-vis des facteurs extérieurs. On peut montrer que dans tous ces cas, les conditions de stabilité des écosystèmes restent invariantes : l'autorégulation par l'interaction de composantes structurales relativement indépendantes, échangeant entre elles l'information sur leur propre état et sur l'état de l'environnement. C'est seulement s'appuyées sur la protection, la restauration, l'imitation des orientations historiquement déterminées et l'intensité des courants informationnels (et énergétiques) que sont possibles aussi bien l'existence durable de la communauté préservée ou nouvellement créée dans un milieu stable que sa faculté de réagir adéquatement à tel ou tel facteur extérieur ne sortant pas des limites de l'optimum d'adaptation.

Pour réaliser des mesures de ce genre, il faut : 1° passer des formulations qualitatives à la création et à l'étude des propriétés de modèles mathématiques et expérimentaux, permettant de faire une prévision durable du comportement des populations naturelles dans les divers types d'action économique ou autre sur leur patrimoine génétique ; 2° embrasser dans des recherches de la génétique des populations les espèces d'animaux et de plantes exploitables les plus importantes pour l'homme ; 3° pour la prise de décisions pratiques, renforcer les contacts entre les collectifs scientifiques et les organisations de production directement rattachées à l'exploitation et à la reproduction des ressources biologiques.

RÉFÉRENCES

[1] *Алтухов Ю. П.* Об иммуногенетическом подходе к проблеме внутривидовой дифференциации рыб. *Усп. совр. генет.* М., Наука, 1969, 2, 161-195.

[2] *Altukhov Yu. P.* Recent physiological, biochemical and immunological studies on the problem of intraspecific differentiation in marine fish. *Cons. Internat. Expl. Mer., Rapp. Proc. Verb.*, 1971, 161, 103-108.

[3] *Алтухов Ю. П.* Локальные стада рыб как генетически стабильные популяционные системы. *Биохим. генет. рыб.* Л., 1973, 43-53.

[4] *Altukhov Yu. P.* Rational management of marine biological resources in the light of population genetics. In : Prop. marine resources Pacific Ocean, Tokyo, Tokai Univ. Press, 1973, 31-39.

[5] *Алтухов Ю. П.* Популяционная генетика рыб. М., Пищев. пр-сть, 1974, 245 с.
[6] *Алтухов Ю. П.* Генетика природных популяций и ресурсы биосферы. *Вестн. АН СССР*, 1975, 10, 37-45.
[7] *Алтухов Ю. П.* Проблемы популяционно-генетической организации вида у рыб. *Ж. общ. Биол.*, 1977, 38, 6, 893-906.
[8] *Altukhov Yu. P.* The stock concept from the viewpoint of population genetics. *Can. J. Fish. Aquat. Sci.*, 1981, 38, 12, 1523-1538.
[9] *Altukhov Yu. P., Nefjedov G. N.* A study of blood serum protein composition by agar-gel electrophoresis in types of red-fish (genus *Sebastes*). *N.-W. Atl. Fish. Comm. Res. Bull.*, 1968, 5, 86-90.
[10] *Алтухов Ю. П., Рычков Ю. Г.* Популяционные системы и их структурные компоненты. Генетическая стабильность и изменчивость. *Ж. общ. Биол.*, 1970, 31, 5, 507-526.
[11] *Алтухов Ю. П., Калабушкин Б. А.* Стабильный полиморфизм в современной и ископаемой популяциях моллюска *Littorina squalida*. *ДАН СССР*, 1974, 215, 6, 1477-1480.
[12] *Алтухов Ю. П., Бернашевская А. Г.* Экспериментальное моделирование динамики генных частот в системе полуизолированных популяций. *ДАН СССР*, 1978, 238, 3, 712-714.
[13] *Алтухов Ю. П., Победоносцева Е. Ю.* Экспериментальное моделирование генетических процессов в подразделенных популяциях. *ДАН СССР*, 1978, 238, 2, 466-469.
[14] *Алтухов Ю. П., Победоносцева Е. Ю.* Исследование биологических особенностей экспериментальной популяционной системы *Drosophila melanogaster*. *Ж. общ. Биол.*, 1979, 40, 3, 368-376.
[15] *Алтухов Ю. П., Победоносцева Е. Ю.* Особенности генетического процесса в экспериментальной популяционной системе *Drosophila melanogaster*. *Ж. общ. Биол.*, 1979. 40, 6, 816-823.
[16] *Алтухов Ю. П., Бернашевская А. Г.* Экспериментальное моделирование генетических процессов в популяционной системе *Drosophila melanogaster*, соответствующей кольцевой ступенчатой модели. II. Стабильность аллельного состава и периодическая зависимость изменчивости частот аллелей от расстояния. *Генетика*, 1981, 17, 6, 1052-1059.
[17] *Altukhov Yu. P., Salmenkova E. A.* Applications of the stock concept to fish populations in the USSR. *Can. J. Fish. Aquat. Sci.*, 1981, 38, 12, 1591-1600.
[18] *Алтухов Ю. П., Апекин В. С., Лиманский В. В.* Основные принципы исследования внутри- и межвидовой дифференциации рыб серологическими методами. М., Пищев. пр-сть, 1964, 53-71.
[19] *Алтухов Ю. П., Салменкова Е. А., Сачко Г. Д.* Дупликация и полиморфизм генов лактатдегидрогеназы у тихоокеанских лососей. *ДАН СССР*, 1970, 195, 3, 711-714.
[20] *Алтухов Ю. П., Угрюмов С. А., Золотарева И. М.* Способ искусственного воспроизводства локальных стад живых организмов. Авт. св. 440131, 1972.
[21] *Алтухов Ю. П., Лиманский В. В., Паюсова А. И., Трувеллер К. А.* Иммуногенетический анализ внутривидовой дифференцировки европейского анчоуса, обитающего в Черном и Азовском морях. I. Группы крови анчоуса и возможный механизм их генного контроля. *Генетика*, 1969, 5, 4, 50-64.
[22] *Алтухов Ю. П., Лиманский В. В., Паюсова А. И., Трувеллер К. А.* Иммуногенетический анализ внутривидовой дифференцировки европейского анчоуса, обитающего в Черном и Азовском морях. II. Элементарные популяции анчоуса и их место в генетико-популяционной структуре вида. *Генетика*, 1969, 5, 5, 81-94.
[23] *Алтухов Ю. П., Салменкова Е. А., Коновалов С. М., Пудовкин А. И.*

Стационарность распределений частот генов лактатдегидрогеназы и фосфоглюкомутазы в системе субпопуляций локального стада рыб (на примере *Oncorhynchus nerka*). I. *Генетика*, 1975, 11, 4, 44-53.

[24] *Алтухов Ю. П., Пудовкин А. И., Салменкова Е. А., Коновалов С. М.* Стационарность распределений частот генов лактатдегидрогеназы и фосфоглюкомутазы в системе субпопуляций локального стада рыб. II. *Генетика*, 1975, 11, 4, 54-62.

[25] *Алтухов Ю. П., Бернашевская А. Г., Милишников А. Н., Новикова Г. А.* Экспериментальное моделирование генетических процессов в популяционной системе *Drosophila melanogaster*, соответствующей кольцевой ступенчатой модели. I. Обоснование подхода и особенности локальной дифференциации частот аллелей α-глицерофосфатдегидрогеназы и эстеразы-6. *Генетика*, 1979, 15, 4, 646-655.

[26] *Алтухов Ю. П., Салменкова Е. А., Рябова Г. Д., Куликова И. И.* Генетическая дифференциация популяций кеты и эффективность некоторых акклиматизационных мероприятий. *Биол. моря*, 1980, 3, 23-38.

[27] *Алтухов Ю. П., Салменкова Е. А., Омельченко В. Т., Сачко Г. Д., Слынько В. И.* О числе мономорфных и полиморфных локусов в популяции кеты — одного из тетраплоидных видов тихоокеанских лососей. *Генетика*, 1972, 8, 2, 251-259.

[28] *Ароншгам А. А., Боркин Л. Я., Пудовкин А. И.* Изоферменты в популяционной и эволюционной генетике. *Генетика изоферментов*. М., Наука, 1977, 199-249.

[29] *Банников А. Г.* Всемирная стратегия охраны природы и охрана животного мира. *Природа*, 1979, 5, 24-28.

[30] *Дягилев С. Е., Маркевич И. Б.* Разновременность созревания горбуши (*Oncorhynchus gorbuscha* Walb.) четных и нечетных лет как основной фактор, определивший различные результаты ее акклиматизации на севере Европейской части СССР. *Вопр. ихтиологии*, 1979, 19, 230-245.

[31] *Dobzhansky Th.* Genetics of natural populations. XXV. A progress report on genetic changes in populations of *Drosophila pseudoobscura* in the American South-West. *Evolution*, 1964, 18, 2, 164-176.

[32] *Дубинин Н. П.* Генетико-автоматические процессы и их влияние на механизмы органической эволюции. *Ж. эксперим. Биол.*, 1981, 7, 463-470.

[33] *Дубинин Н. П.* Общая генетика. М., Наука, 1976, 590 с.

[34] *Дубинин Н. П., Ромашов Д. Д.* Генетическое строение вида и его эволюция. *Биол. Ж.*, 1932, 1, 5-6, 52-95.

[35] *Fisher R. A.* The genetical theory of natural selection. Oxford, Clarendon Press, 1930.

[36] *Haldane J. B. S.* The causes of evolution. N. Y.-L., Harper, 1932.

[37] *Кирпичников В. С.* Генетические основы селекции рыб. Л., Наука, 1979, 392 с.

[38] *Коновалов С. М.* Структура изолята нерки озера Азабачьего. *Ж. общ. Биол.*, 1972, 33, 6, 668-682.

[39] *Коновалов С. М.* Популяционная биология тихоокеанских лососей. Л., Наука, 1980, 237 с.

[40] *Коновалов С. М., Шапиро А. П., Лейбович Г. Е.* Эксплуатация биологических ресурсов в связи с пространственной структурой вида. *Биол. моря*, 1975, 6, 26-36.

[41] *Малиновский А. А., Мина М. В.* Рецензия на кн. Ю. П. Алтухова «Популяционная генетика рыб». *Ж. общ. Биол.*, 1976, 37, 5, 788-790.

[42] *Мина М. В.* О популяционной структуре вида у рыб. *Ж. общ. Биол.*, 1978, 39, 3, 453-460.

[43] *Ricker W. E.* Hereditary and environmental factors affecting certain salmonid populations. The stock concept of Pacific salmon. H. R. McMillan Lect. in Fisheries, Univ. Brit. Columbia, 1972, 19-160.

[44] *Шмальгаузен И. И.* Организм как целое в индивидуальном и историческом развитии. М.-Л., Изд-во АН СССР, 1938, 144 с.
[45] *Sokolov V.* The biosphere reserve concept in the USSR. *AMBIO*, 1981, 10, 2-3, 97-101.
[46] *Старобогатов Я. И.* Рецензия на кн. Ю. П. Алтухова « Популяционная генетика рыб ». *Ж. общ. Биол.*, 1975, 36, 4, 626-628.
[47] *Тимофеев-Ресовский Н. В., Воронцов Н. Н., Яблоков А. В.* Краткий очерк теории эволюции. М., Наука, 1977, 301 с.
[48] *Четвериков С. С.* О некоторых моментах эволюционного процесса с точки зрения современной генетики. *Ж. эксперим. биол.*, 1926, 1, 3-54.
[49] *Чунихин С. П.* Экологические потрясения. *Химия и жизнь*, 1979, 2, 32-35.
[50] *Wright S.* Evolution in Mendelian populations. *Genetics*, 1931, 16, 97-159.
[51] *Wright S.* Evolution and the genetics of populations. Vol. 2. The theory of gene frequencies. Chicago, Univ. Chicago Press, 1969, 511 p.

VARIABILITÉ DANS UNE POPULATION ACTUELLE ET DANS UNE POPULATION DE L'HOLOGÈNE MOYEN DE *LITTORINA SQUALIDA*

B. Kalabouchkine

Dans les recherches de la génétique évolutive, la population est considérée comme la « structure évolutive élémentaire de l'espèce » [26]. La complexité et la diversité de l'organisation biologique des populations sont largement mises en lumière dans les travaux d'E. Mayr [22], de N. Timoféev-Ressovski [26] et de nombreux autres auteurs. Sous le rapport génétique, ces structures élémentaires ne sont pas non plus homogènes, ce sont des ensembles plus ou moins hétérogènes. Ceci a permis à Y. Altoukhov et à Y. Rytchkov [4] de montrer la stabilité de ces systèmes au fil des générations et de distinguer deux niveaux qualitatifs de leur organisation, le premier étant les systèmes de populations, le second, les composantes structurales.

Dans la majorité des cas, les modifications survenant dans les populations au cours d'une longue période restent inaccessibles à l'observation. Cependant, on peut parfois conclure à leur état en confrontant les paramètres génétiques à des modèles mathématiques connus [3, 24] ou en comparant directement les populations de périodes différentes. Par cette approche, il est possible d'établir un rapport entre le degré de modifications génétiques et l'histoire naturelle de la population étudiée, ainsi que la durée de son habitation dans l'aire envisagée. Les travaux de G. Currey et d'A. Cain [8] sur les rapports entre les phénotypes dans les populations de *Cepea nemoralis* et *C. hortensis* d'Angleterre méridionale au cours de 6500 ans, ceux de Y. Rytchkov et A. Movsessian [23] sur la répartition des anomalies du crâne dans les populations mongoloïdes néolithique et actuelle d'Asie septentrionale sont des exemples de recherches de ce genre.

On a déjà montré [5] la stabilité de la fréquence du gène dans une population de *Littorina squalida* habitant la lagune de Boussé au cours des dernières 4500-5500 années pour ce qui concerne la coloration des coquilles des mollusques. Le présent travail se propose de mettre en évidence les facteurs déterminant la structure génétique de la population et les mécanismes de sa stabilisation.

© Издательство Наука, 1976.

MATÉRIEL ET MÉTHODE

Le mollusque *Littorina squalida* Broderip et Sowerby, 1829 (classe des gastropodes, ordre des discopodes, famille des littorinidés) est une espèce boréale largement répandue dans le Pacifique Nord. La littorine vit dans la mer du Japon sur les côtes de l'île de Hokkaidô et des Kouriles, dans la mer d'Okhotsk et de Béring sur les côtes des Aléoutiennes, des îles de Pribylov, du Kamtchatka et de l'Alaska. On la rencontre depuis les horizons supérieurs du littoral jusqu'à une profondeur de 10 à 15 m (golfe Avatchinski), elle se loge ordinairement dans l'horizon inférieur du littoral [12, 13]. C'est une espèce bisexuée à fécondation interne. Au cours de la reproduction, la littorine forme des capsules planctoniques d'œufs qui permettent la propagation de l'espèce. Une fois fixés, les mollusques développés n'effectuent pas de migrations importantes.

Le matériel a été recueilli en 1969-1973 dans la lagune de Boussé (golfe Anivski, île de Sakhaline) ; le matériel subfossile, dans le lit du fleuve sur la côte orientale de la lagune.

La population de la lagune est représentée dans les trois horizons du littoral : supérieur, moyen et inférieur. Dans la lagune de 7 à 7,5 km de diamètre communiquant avec le golfe Anivski par un détroit d'environ 100 m de large, les deux premiers horizons sont représentés par les régions côtières de 15 à 400 m de large, un banc d'huîtres et les prairies de zostères aux eaux les plus basses. L'horizon inférieur embrasse les portions des prairies de zostères qui restent à marée basse sous une couche d'eau de 60 à 80 cm. La population actuelle est caractérisée d'après 40 échantillons (environ 4000 individus). Chaque échantillon a été choisi sur une superficie de 1,5 à 2 m^2 au maximum au fond de la lagune et dans les couches d'algues. Le nombre des épreuves pour chaque zone littorale est à peu près proportionnel à l'aire de ces zones et à la densité des subpopulations y habitant. Le matériel subfossile (6 échantillons, 479 individus) a été pris sur des secteurs limités et ordinairement dans une couche de 30 à 50 cm d'épaisseur et sur une surface de 1 m^2 environ.

Afin de déterminer l'âge et la reconstruction paléogéographique, arrêtons-nous à l'histoire de la formation de la lagune. Les côtes et les langues de terre qui la séparent du golfe Anivski sont formées par deux niveaux de terrasses d'accumulation correspondant à deux stades de formation de la lagune (fig. 1). La hauteur de la première terrasse est à peu près de 1,5 à 2 m ; celle de la seconde, de 5 à 6 m. Il faut signaler l'accroissement prononcé de la hauteur des terrasses vers les régions de l'estuaire où elles atteignent respectivement 5, 10 et 12 m et la baisse de la hauteur sur les côtes intérieures. La formation des terrasses, tout au moins aux étapes finales, avait lieu dans la zone du supralittoral, ce qui explique leurs variations de hauteur. Au niveau de la seconde terrasse correspondent des alluvions riches

en faune de mollusques gastropodes et bivalves, découverts dans le lit du fleuve sur la côte orientale de la lagune. Parmi les mollusques bivalves on rencontre des espèces dont la limite nord d'habitation est actuellement déplacée vers le sud. C'est ainsi que pour les espèces *Arca boucardi* et *Gastrana contalulata*, c'est la côte nord de Hokkaidō, et pour l'espèce *Trapezium liratum*, le golfe de Pierre le Grand [13]. Des trouvailles de faune thermophile permettent de dater ces alluvions du temps de l'optimum climatique de l'holocène moyen

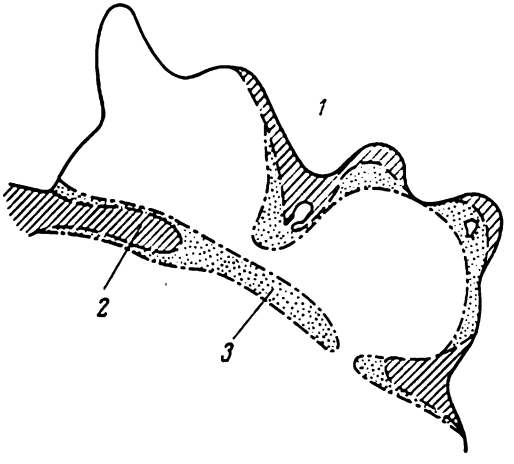


Fig. 1. Reconstruction des lignes côtières dans la lagune de Boussé:
1 — holocène inférieur-pléistocène supérieur, *2* — holocène moyen, *3* — holocène supérieur

fixé selon le complexe d'algues diatomées [14] et les données palynologiques [1, 17]. Suivant les estimations obtenues par N. Khotinski [16], l'optimum climatique de l'holocène moyen à Sakhaline est daté de 7000 à 9000 ans. Cependant, compte tenu du caractère des alluvions qui atteste leur formation aux étapes terminales de la transgression ayant eu lieu il y a de 4000 à 7000 ans [15], il semble plus exact de dater les alluvions de l'holocène moyen de la lagune de Boussé de 4500 à 5500 ans. Comme le montre la figure 1, à mesure que la ligne des côtes se développait, l'aire du bassin intérieur de la lagune se réduisait et son degré d'isolement du golfe Anivski augmentait.

Les conditions d'enfouissement et la conservation permettent de considérer les prélèvements de matériel subfossile comme des échantillons de la population de l'holocène moyen de Protoboussé.

La netteté des structures géomorphologiques et les données relatives à la répartition des espèces dans les alluvions, dont la datation écologique est bien connue dans la lagune actuelle, permettent de distinguer des zones littorales dans le golfe de l'holocène moyen.

Comme nous le montrerons plus loin, la structure de la population actuelle est déterminée essentiellement par les différences observées en comparant les subpopulations se rapportant aux horizons supérieur, moyen et inférieur du littoral et relativement homogènes dans les limites de chacun de ces horizons. C'est pourquoi les échantillons de matériel subfossile de toutes les zones littorales reflètent, apparemment, d'une façon assez juste l'état de la population ancestrale, bien qu'une de ses parties seulement soit étudiée. Notons les différences d'estimations des paramètres dans les populations actuelle

Fig. 2. Polymorphisme de la coloration des coquilles de *L. squalida*: phénotypes: *1* — A, *2* — AB, *3* — B

et de l'holocène moyen. Si, dans le premier cas, on analyse trois ou quatre générations successives, dans la population ancestrale on étudie un mélange de dizaines ou de centaines de générations.

D'après la coloration et le dessin des coquilles de *L. squalida* on peut isoler trois types possibles: 1° des coquilles claires sans raies; 2° des coquilles rayées (quelquefois les traces de raies ne sont visibles que sur la face interne de la dernière spire) variant du jaune grisâtre à des couleurs foncées et toujours avec une raie claire sur le côté inférieur de la dernière spire; 3° des coquilles rayées noires ou très foncées avec une raie claire étroite sur la face interne de la dernière spire (fig. 2). Ces types de dessin de la coquille sont suivis dans tous les groupes d'âge du matériel recueilli par l'auteur dans les îles de Sakhaline, de Kunachir, ainsi que dans les collections examinées de l'Institut de Zoologie de l'Académie des Sciences de l'U.R.S.S., recueillies dans le golfe Possiet, sur la côte de Magadan et dans la presqu'île du Kamtchatka. On avait postulé antérieurement l'hypothèse du contrôle génique à dominance incomplète dans l'interprétation suivante: type 1 — homozygote AA, type 2 — hétérozygote AB, type 3 — homozygote BB [5].

RÉSULTATS

L'étude chorologique de la population de la lagune de Boussé a mis en évidence la structure nettement prononcée de subpopulations avec les fréquences géniques: dans les régions côtières, le littoral supérieur et, en partie, moyen $pA = 0{,}045$ (15 échantillons, 1006

individus); sur le banc d'huîtres, le littoral supérieur et moyen, $pA = 0{,}169$ (5 échantillons, 478 individus); dans les prairies de zostères, le littoral moyen et inférieur $pA = 0{,}449$ (20 échantillons, 2334 individus). Comme le montrent ces données, les fréquences géniques dépendent pour beaucoup de la localisation des subpopulations. De plus, les variations de fréquences des génotypes sont en corrélation avec le gradient d'environnement. Dans le cas d'un changement graduel des zones littorales, on observe une variabilité clinale progressive et dans le cas où elles sont nettement délimitées, on note de brusques différences. L'analyse des fréquences des génotypes dans les générations met en évidence une réduction marquée de la part des homozygotes dans les groupes d'âge avancé (*Tableau 1*).

Tableau 1

Dynamique des fréquences des génotypes dans les groupes d'âge de trois subpopulations de la lagune

Groupe d'âge	Régions côtières			Banc d'huîtres			Prairies de zostères					
							littoral moyen			littoral inférieur		
	A	AB	B	A	AB	B	A	AB	B	A	AB	B
t_0				0,010	0,270	0,720	0,013	0,573	0,414	0,594	0,380	0,026
t_1		0,085	0,915	0,016	0,365	0,619	0,008	0,868	0,124	0,251	0,731	0,018
t_2		0,121	0,879	0,012	0,360	0,689	0,001	0,958	0,041	0,000	0,989	0,011
t_3		0,185	0,815	0,000	0,670	0,330	0,000	1	0,000			
t_4		0,250	0,750				0,000	1	0,000			

Avant de passer à l'examen des facteurs d'hétérogénéité spatiale et de dynamique des fréquences géniques dans les populations, il faut remarquer que la population de la lagune est un isolat dans l'analyse duquel on peut négliger l'apport de gènes des populations voisines. La population la plus proche d'individus de cette espèce avec une densité et un effectif beaucoup moindres est éloignée de l'estuaire de la lagune de 12 à 15 km. Les particularités de la biologie de la reproduction des mollusques [18] permettent de considérer l'isolat comme une population panmictique. Dans les conditions de la lagune les mollusques parviennent à la maturité sexuelle au cours de la deuxième année de leur vie. Cette conclusion est tirée des résultats de l'ouverture et de la comparaison des dimensions des coquilles des mollusques de différents groupes d'âge avec les données d'A. Alimov et d'A. Golikov [2].

Tout en restant dans le cadre de l'hypothèse suggérée du contrôle génique, il faut considérer les données relatives à la variation de la composition génotypique des générations et la présence d'une structure dans la population comme le témoignage d'une survie diffé-

rentielle des individus appartenant à des génotypes différents. Dans l'ensemble, l'intensité de la sélection peut dépendre considérablement de la densité de la population, des variations saisonnières de l'environnement et des conditions microécologiques. Sans disposer de données pour analyser ces facteurs, nous essayerons d'estimer les valeurs moyennes de la sélection en trois endroits de la lagune à différentes périodes.

Admettons que l'intensité de la sélection ne dépende pas de l'âge de la génération et soit constante dans le temps, posons les variations de fréquences des génotypes dans une série de générations successives égales aux variations de chaque génération dans le temps, nous pourrons alors évaluer la probabilité de mort des génotypes durant l'année à 1 — $V = S$ où V est la probabilité de survie durant l'année et S un coefficient de sélection. Définissons V comme le rapport $V = n_{i+1}/n_i$ où n_i, n_{i+1} sont les fréquences relatives des génotypes dans les groupes d'âge t_i et t_{i+1}. Dans ce cas, la probabilité de survie des génotypes au cours de l'année est égale à

$$V_{11} = \frac{n^{11}_{i+1}}{n^{11}_{i}}, \quad V_{12} = \frac{n^{12}_{i+1}}{n^{12}_{i}}, \quad V_{22} = \frac{n^{22}_{i+1}}{n^{22}_{i}}.$$

Pour trouver la probabilité de survie des génotypes dans chaque secteur, adressons-nous aux données du *Tableau 1* dans lequel sont présentées les fréquences moyennes des génotypes dans les groupes d'âge de trois subpopulations de la lagune. A partir du rapport des génotypes on obtient les valeurs pour chaque paire de générations voisines et en faisant la moyenne sur ces valeurs, on trouve leurs moyennes arithmétiques.

Afin de vérifier la correspondance des estimations obtenues avec les variations réelles de la composition génotypique des générations, servons-nous de la distribution des génotypes dans le groupe de mollusques de l'année et calculons, en utilisant les valeurs de $\overline{V}_{ij}$ données plus haut, les effectifs théoriques des génotypes dans les générations t_1, t_2, t_3 et t_4 (*Tableau 2*). Comme le montrent les valeurs mentionnées de χ^2, les valeurs calculées reflètent de façon satisfaisante la dynamique des variations dans les régions côtières et dans les prairies de zostères. Quant à la subpopulation du banc d'huîtres, on observe un écart significatif entre les effectifs correspondants des génotypes ($\chi^2 = 20{,}32 \gg \chi^2_{05} = 6{,}63$). Ceci s'explique non seulement par l'insuffisance de matériel, mais aussi par une hétérogénéité importante de la subpopulation, déterminée aussi bien par sa répartition sur une aire fort réduite dont les conditions subécologiques varient brusquement, que par une grande influence des migrations et de l'apport venant des zones avoisinantes du littoral inférieur et moyen.

Tableau 2

Effectifs moyens observés et théoriques des génotypes dans les groupes d'âge de trois subpopulations de la lagune

Groupe d'âge	Régions côtières			Banc d'huîtres			Prairies de zostères					
							littoral moyen			littoral inférieur		
	A	AB	B	A	AB	B	A	AB	B	A	AB	B
t_1				1	23	39	1,90	219,6	31,50	19,86	57,75	1,39
				(0,44)	(25,07)	(37,48)	(1,16)	(209,51)	(42,23)	(21,89)	(56,04)	(1,07)
t_2		37,33	271,67	1	25	57	0,95	652,67	27,38	0,000	28,67	0,33
		(38,66)	(270,34)	(0,42)	(44,99)	(37,60)	(0,82)	(648,8)	(36,36)	(2,56)	(26,3)	(0,14)
t_3		10,38	45,62	0	12	6	0,00	197,84	2,16			
		(10,10)	(45,90)	(0,08)	(12,20)	(5,72)	0,00	(196,84)	(3,10)			
t_4		7,5	22,5				0,00	27,00	0,00			
		(7,58)	(22,42)				(0,00)	(26,88)	(0,12)			
Valeurs de χ^2	$0,07 < \chi^2_{05} = 3,84$			$20,32 \gg \chi^2_{01} = 6,63$			$\chi^2_{05} = 5,99 < 6,8 < \chi^2_{01} = 9,21$					

Remarque. Entre parenthèses on indique les effectifs théoriques. Les valeurs de χ^2 sont données aux niveaux de signification 0,05 et 0,01.

Tableau 3

Adaptabilité relative des génotypes aux stades ontogénétiques précoces (L) et tardifs (V), valeurs globales (W) de l'adaptabilité dans les différents secteurs de la lagune

Adaptabilité relative	Régions côtières			Banc d'huîtres			Prairies de zostères					
							littoral moyen			littoral inférieur		
	A	AB	B	A	AB	B	A	AB	B	A	AB	B
L	0	0,06	1	0,05	0,37	1	0,09	1	0,74	1	0,16	0,01
V		1	0,65	0,53	0,37	0,56	0,25	1	0,28	0,25	1	0,28
W	0	0,09	1	0,05	0,66	1	0,02	1	0,21	1	0,65	0,12

Les données mentionnées plus haut sur le moment de la maturation sexuelle permettent de rattacher les mollusques d'un an à la partie reproductrice de la population, ce qui permet, à son tour, de considérer les grandeurs $\bar{V}_{ii}$ comme la probabilité d'atteindre l'âge de la reproduction ou comme l'adaptabilité relative des génotypes aux stades ontogénétiques avancés. En pondérant par l'effectif la contribution de chacune des subpopulations à la partie reproductrice de la population [21], trouvons, compte tenu de la loi de Hardy-

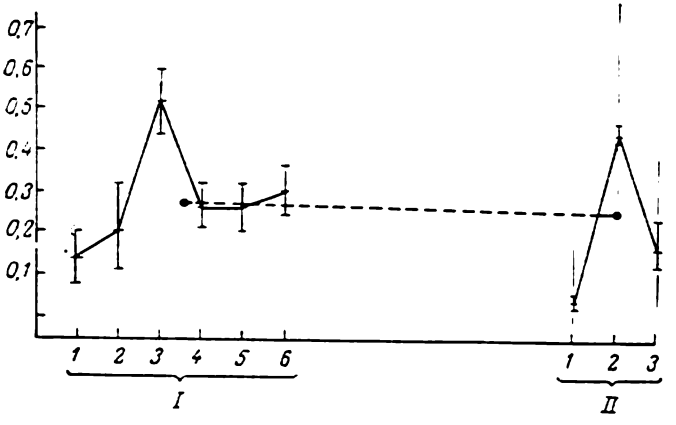


Fig. 3. Variabilité dans les populations et entre populations dans une population actuelle et une population de l'holocène moyen de *L. squalida* de la lagune de Boussé:

I — population de l'holocène moyen: fréquences géniques et intervalles de confiance au niveau 95 % dans les échantillons: *1, 2* — des régions côtières; *3* — des secteurs centraux; *4-6* — sur le banc d'huîtres; *II* — population actuelle: fréquences géniques, intervalles de confiance au niveau 95 % et ampleur de la variabilité des fréquences géniques dans les échantillons de subpopulations: *1* — des régions côtières; *2* — des secteurs centraux; *3* —sur le banc d'huîtres.

La ligne pointillée réunit les moyennes des fréquences géniques pour l'ensemble de la population

Weinberg, les fréquences théoriques des génotypes dans la descendance. A la condition d'une dispersion aléatoire et du dépôt des capsules planctoniques sur toutes les parties de la lagune, on peut attendre, dans le groupe des mollusques de l'année, un rapport de génotypes typique de la population dans son ensemble et égal à 0,114, 0,448 et 0,438 respectivement pour AA, AB et BB. A partir du rapport des fréquences des génotypes observées dans le groupe des mollusques de l'année aux fréquences calculées, on peut obtenir les valeurs de l'adaptabilité des génotypes et apprécier respectivement l'intensité de la sélection agissant dans un bref laps de temps, probablement après le dépôt des capsules planctoniques aux stades les plus précoces du développement. Les valeurs de l'adaptabilité trouvées sont présentées au *Tableau 3*. Comme le montrent ces données, la direction et

l'intensité de la sélection dépendent de la localisation des subpopulations. Les valeurs globales de l'adaptabilité des génotypes dans trois secteurs étudiés de la lagune sont égales au produit des adaptabilités des génotypes aux stades précoces et tardifs de l'ontogénèse et sont mentionnées au *Tableau 3*.

La distribution spatiale de fréquences géniques dans la population de l'holocène moyen s'avère très proche de celle observée actuellement. On parvient aussi à découvrir dans la population ancestrale des zones de basses fréquences se rapportant aux régions côtières (échantillons 1 et 2), à la région du banc d'huîtres (échantillons 4, 5 et 6) et des fréquences relativement élevées dans les parties centrales (échantillon 3) (fig. 3). A l'époque, la population ancestrale éprouvait la pression de sélection conduisant à la formation de la structure. La comparaison de la distribution totale des génotypes de toutes les générations de la population actuelle et de celle de l'holocène moyen met en évidence une différence significative entre ces populations: $\chi^2 = 45{,}2 \gg \chi^2_{.01} = 6{,}63$. La même comparaison, concernant maintenant le groupe des mollusques de l'année de la population actuelle, montre leur grande ressemblance: $\chi^2_{.05} = 3{,}84 < \chi^2 = 6{,}44 < \chi^2_{.01} = 6{,}63$. A la différence de la population actuelle, on ne découvre pas dans la population ancestrale un excès d'hétérozygotes dans les échantillons. Probablement que dans les conditions du golfe Protoboussé, il n'y avait pas de différences d'adaptabilité des génotypes aux stades ontogénétiques tardifs. Sous ce rapport, la population ancestrale ressemble beaucoup à la population des côtes ouvertes de la mer d'Okhotsk sur lesquelles on n'observe pas non plus d'excès d'hétérozygotes dans les échantillons [5] et où les estimations des fréquences géniques dans les groupes d'âge sont très proches les unes des autres.

DISCUSSION

A cause des difficultés à réaliser des croisements, l'hypothèse proposée de l'hérédité des variantes de coloration des coquilles de *L. squalida* est acceptée sans démonstration. Dans ce cas, on pourrait s'en tenir au niveau de l'analyse phénétique [26]. Cependant, certaines prémisses adoptées dans le travail ont permis, semble-t-il, d'interpréter plus largement les données obtenues.

Partant de la supposition de la dispersion aléatoire et du dépôt des capsules planctoniques dans le bassin intérieur de la lagune, on a trouvé les coefficients de sélection agissant dans un bref laps de temps aux stades ontogénétiques précoces. La variation de fréquences des génotypes dans une suite de générations a permis de déterminer la probabilité de leur mort au cours de l'année. En se fondant sur ces données, on a trouvé les valeurs globales de l'adaptabilité relative des génotypes: dans le littoral supérieur et moyen des régions cô-

tières $W_{AA} = 0$; $W_{AB} = 0,09$; $W_{BB} = 1$; sur le banc d'huîtres $W_{AA} = 0,05$; $W_{AB} = 0,66$; $W_{BB} = 1$; dans le littoral moyen des prairies de zostères $W_{AA} = 0,02$; $W_{AB} = 1$; $W_{BB} = 0,21$; dans le littoral inférieur des prairies de zostères $W_{AA} = 1$; $W_{AB} = 0,65$; $W_{BB} = 0,12$. Les valeurs obtenues des coefficients de sélection doivent être considérées comme des estimations qualitatives indiquant l'orientation de la sélection. Il est probable que des valeurs aussi faibles de l'adaptabilité relative ne soient pas une exception et puissent être rencontrées dans des populations de différentes espèces, surtout celles d'une grande fécondité. Ainsi, par exemple, S. Guerchenson [11] a trouvé des coefficients de sélection des hamsters noirs oscillant entre 0,86 et 0,96. Les travaux de G. Antonovics [6] et de R. Berry [7] groupent en tableau les données de nombreux auteurs sur les valeurs de la sélection dans des populations de plantes et d'animaux, chez certaines populations elles atteignent 0,6-0,8. Il est encore impossible d'expliquer de façon satisfaisante l'avantage de tels ou tels génotypes dans différentes conditions écologiques. De toute évidence, comme chez les mollusques terrestres [9, 19], ils sont en corrélation avec toutes sortes d'autres caractères. Sans nous arrêter sur toute la diversité des conditions déterminant la structure génétique des populations, signalons seulement les conditions, les plus importantes pour notre ligne d'examen, liées à la nature adaptative du polymorphisme. On sait que les conditions changeant, les fréquences des gènes ou des génotypes dans les populations peuvent varier sous la pression de sélection [8, 10]. Dans d'autres cas, la sélection déterminée par l'hétérogénéité spatiotemporelle du milieu peut être un facteur de stabilisation [22]. Comme le montrent nos données, il se produit, dans la population de la lagune, une sélection aux stades ontogénétiques précoces ayant des directions différentes dans les sous-niches écologiques et une sélection au profit des hétérozygotes dans les stades tardifs. Notons que, dans la période de temps étudiée, l'aire d'habitation de la population ne restait pas constante, mais augmentait à mesure que des bas-fonds se formaient dans les régions côtières et centrales de la lagune. L'exploitation de nouveaux secteurs a entraîné des différences dans la distribution de fréquences des génotypes entre la population actuelle et la population de l'holocène moyen. De plus, la pression croissante de sélection aux stades précoces a influé sur la distribution des génotypes dans une moindre mesure que la nouvelle forme de sélection, la superdominance créant un excès d'hétérozygotes dans la population actuelle. Les deux types de sélection sont des facteurs stabilisants [20, 25] si bien que l'examen de la population utilisant les paramètres des fréquences géniques ne révèle pas la moindre variation de fréquence du gène durant une période d'environ 2000 générations successives.

RÉFÉRENCES

[1] *Александров С. Н.* Остров Сахалин. М., Наука, 1973, 1-182.
[2] *Алимов А. Ф. Голиков А. Н.* Некоторые закономерности соотношения между размерами и весом у моллюсков. *Зоол. Ж.*, 1974, 53, 4, 517-530.
[3] *Алтухов Ю. П.* Популяционная генетика рыб. М., Пищевая промышленность, 1974, 1-247.
[4] *Алтухов Ю. П., Рычков Ю. Г.* Популяционные системы и их структурные компоненты. Генетическая стабильность и изменчивость. *Ж. общ. Биол.*, 1970, 31, 5, 507-526.
[5] *Алтухов Ю. П., Калабушкин Б. А.* Стабильный полиморфизм в современной и ископаемой популяциях моллюсков *Littorina squalida*. *Докл. АН СССР*, 1974, 215, 6, 1477-1480.
[6] *Antonovics G.* The effects of a heterogeneous environment on the genetics of natural populations. *Am. Scien.*, 1973, 59, 593-599.
[7] *Berry R. S.* The II-nd Symposium of British Ecological Society, Blackwell Scientific Publications, Oxford, 1973, 177-206.
[8] *Currey G. D., Cain A. G.* Studies on *Cepaea*. IV. Climate and selection of banding morphs in *Cepaea* from the climatic optimum to the present day. *Philos. Tr. Roy. Soc. London*, 1968, 483-494.
[9] *Ehrlich P., Holm R.* The Process of evolution. N. Y., 1963.
[10] *Ford E. B.* Ecological genetics. N. Y., 1964, 1-268.
[11] *Гершензон С. М.* Роль естественного отбора в распространении и динамике меланизма у хомяков. *Ж. общ. Биол.*, 1946, 7, 2, 97-127.
[12] *Голиков А. Н., Кусакин О. Г.* Фауна и экология брюхоногих переднежаберных моллюсков литорали Курильских островов. *Исслед. дальневост. морей*, 1962, 8, 248-346.
[13] *Голиков А. Н., Скарлато О. Л.*. Моллюски залива Посьет (Японское море) и их экология. *Тр. Зоол. ин-та АН СССР*, 1967, 42, 5-154.
[14] *Жузе А. Н.* К палеогеографии Охотского моря. *Изв. АН СССР*, Сер. геогр., 1959, 2, 12.
[15] *Каплин Н. А.* Развитие береговой зоны морей и океанов в послеледниковое время. Автореф. канд. диссерт. М., Ин-т океанологии, 1970.
[16] *Хотинский Н. А. Шулия К. С.* Голоцен северной Евразии. Тез. докл. III Международной палинолог. конф. Новосибирск, 1971.
[17] *Худяков Г. И., Денисова Е. П., Короткий А. М., Кулаков А. П., Пикунова Р. И., Чернобровкин Е. И.*, Юг Дальнего Востока. М., Наука, 1972, 1-422.
[18] *Koyma J.* On the floating egg capsule of periwinkles, *Littorina squalida* and *Nodilittorina pyramidalis*. *The Venus*, 1958, 19, 3-4, 223-237.
[19] *Lamotte M.* Polymorphism of natural populations of *Cepaea nemoralis*. *Cold. Spring. Harbor Symp. Quant. Biol.*, 1959, 24, 65-86.
[20] *Levene H.* Genetic equlibrium when more than one ecological niche available. *Am. Nat.*, 1953, 87, 311-313.
[21] *Li C. C.* Population genetics. Univ. Press Chicago, 1955, 346.
[22] *Mayr E.* Animal species and evolution. London, 1965.
[23] *Рычков Ю. Г., Мовсесян А. А.* Генетико-антропологический анализ распределения аномалий черепа у монголоидов Сибири в связи с проблемой их происхождения. *Бюл. МОИП.*, 1972, 43, 114-132.
[24] *Рычков Ю. Г., Русакова О. А., Раппопорт М. П., Шереметьева В. А.* Факторы генетической дифференциации популяционной системы коренного населения Северной Азии. Сообщ. I. Определение интегрирующих значений адаптивной ценности генотипов групп крови АВО в коренном населении Сибири. *Генетика*, 1973, 9, 2, 136-142.
[25] *Тимофеев-Ресовский П. В., Свирежев Ю. М.* Об адаптационном полиморфизме в популяциях *Adalia bipunctata*. *Пробл. кибернет.*, 1965, 16, 137-146.
[26] *Тимофеев-Ресовский Н. В., Воронцов Н. Н., Яблоков А. В.* Краткий очерк теории эволюции. М. Наука, 1969.

SIMULATION EXPÉRIMENTALE DES PROCESSUS GÉNÉTIQUES DANS UN SYSTÈME DE POPULATIONS DE *DROSOPHILA MELANOGASTER* CORRESPONDANT AU MODÈLE ÉCHELONNÉ CIRCULAIRE

Communication I. Justification de l'approche et particularités de la différenciation locale des fréquences d'allèles de l'α-phosphoglycérate-déshydrogénase et de l'estérase-6

Y. Altoukhov, A. Bernachevskaïa, A. Milichnikov et T. Novikova

Dans la théorie de la génétique des populations, on estime que ce ne sont ni une grande population panmictique ni une petite population isolée qui sont adaptées pour le mieux à la transformation évolutive, mais un groupement divisé en une série de sous-unités en partie isolées et rattachées entre elles par une interaction de migration et soumises à la dérive génétique [14, 15, 20, 31].

Il a été démontré mathématiquement [14-16, 20, 31] que dans une population subdivisée, comparée à une population panmictique de grandeur confrontable, il y a beaucoup plus de diversité génétique lui permettant de réagir avec davantage d'efficacité aux variations de l'environnement et de modifier en conséquence sa structure génotypique.

Cependant, cette conclusion théorique n'avait été confirmée ni par des recherches en champ ni par l'expérimentation. Relativement peu, on a découvert dans les recherches sur des populations naturelles subdivisées une qualité systémique importante qui leur est propre et qui ne ressort évidemment pas des propriétés de leurs composantes structurales: la stabilité génétique temporospatiale. Il s'est avéré que le processus génétique dans un système de populations se déroule suivant le type stationnaire déterminé par les pressions mutuellement équilibrées que les facteurs aléatoires et systématiques de l'évolution exercent sur sa structure interne [10].

Il est désormais évident que dans la mesure dans laquelle la stabilité génétique du système de populations est en corrélation avec sa diversité interne, les transformations évolutives du système en tant que tout ne seraient guère possibles lors de changements de l'environnement ne sortant pas des limites de son optimum d'adaptation historiquement constitué. Dans les conditions d'un milieu variant normalement, la fréquence génique moyenne et, le maximum d'adaptation étant atteint, la dispersion dans le système doivent

© Издательство Наука, 1979.

rester invariantes par rapport à toutes les transformations microévolutives possibles, différemment dirigées et pour cette raison s'équilibrant mutuellement [3].

Il a donc paru intéressant de mettre au point une approche de la simulation expérimentale des phénomènes génétiques dans un système de populations. D'une part, cela pourrait avoir de l'importance pour une argumentation complémentaire des idées développées et, d'autre part, contribuer à justifier la valeur prédictive de certains modèles cybernétiques indispensables pour résoudre le problème de l'utilisation rationnelle des ressources biologiques. Ce genre de recherches que nous avons entreprises en 1972 [4, 7] ont confirmé les conclusions faites au cours de l'étude des systèmes de populations dans la nature. Par la publication présente, nous entreprenons une série de travaux décrivant en détail les résultats obtenus et l'expérience acquise. La communication I examinera les particularités méthodiques de l'approche et les données concernant la différenciation locale stable des fréquences géniques dans un système de populations partiellement isolées de *D. melanogaster*, organisé selon le type du modèle échelonné circulaire.

MATÉRIEL ET MÉTHODE

En 1974-1977, on a étudié en laboratoire un système de subpopulations de *D. melanogaster* dont la structure de migration correspondait à un modèle échelonné fermé unidimensionnel (stepping-stone) (fig. 1). Du point de vue théorique, le modèle échelonné a été étudié assez en détail par de nombreux auteurs [17, 18, 21]. Cependant, nous avons été les premiers à effectuer des recherches expérimentales de ce genre.

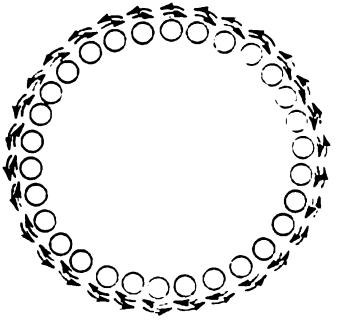


Fig. 1. Particularités de la structure d'un système expérimental de populations en partie isolées de *D. melanogaster*: Les flèches indiquent les directions de la migration éventuelle

L'expérience a eu lieu dans une boîte à population en plexiglas divisée en 30 compartiments. Les compartiments communiquaient entre eux par des tubes placés dans les orifices de migration par lesquels les individus d'une subpopulation se déplaçaient dans une autre. Chaque subpopulation pouvait échanger des individus seulement avec deux subpopulations voisines.

Pour le choix de la longueur et du diamètre des tubes de migration déterminant la valeur du coefficient de migration, une série de 20 expériences a été faite. On a finalement choisi des tubes

de migration de 35 mm de long et de 4 mm de diamètre, ce qui assurait une valeur moyenne du coefficient de migration de 0,027 ± 0,007.

On a utilisé pour l'expérience deux lignées hétérogènes sous le rapport génétique de *D. melanogaster* issues de populations naturelles de la Crimée et du Caucase du Nord. La première lignée a été tirée directement de la population naturelle, la seconde a été entretenue pendant 3 ans dans les conditions de laboratoire. En qualité de marqueurs génétiques on a utilisé deux enzymes séparables par électrophorèse, codées par des loci autosomiques non couplés: Est-6

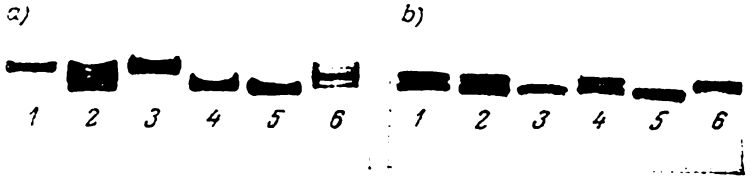


Fig. 2. Photographies des variantes électrophorétiques d'α-PGDH (*a*) et d'Est-6 (*b*):

a — *1* et *3* — homozygotes SS, *4* et *5* — homozygotes FF, *2* et *6* — hétérozygotes FS; *b* — *1*, *2* et *4* — hétérozygotes FS, *3* et *5* — homozygotes SS, *6* — homozygote FF. L'anode est en bas; la cathode, en haut. Electrophorèse sur gel de polyacrylamide suivie de coloration histochimique

(chromosome 3) et α-PGDH (chromosome 2). Pour identifier les enzymes, on a eu recours à l'électrophorèse sur gel de polyacrylamide suivie de coloration histochimique afin de mettre en évidence l'activité enzymatique. La méthode de l'électrophorèse a été exposeé en détail antérieurement [8]. La figure 2 montre les photographies des variantes électrophorétiques de protéines.

5 lignées homozygotes pour les allèles de chaque locus ont été tirées de chaque lignée originelle de *D. melanogaster*. Puis, nous avons obtenu par croisement des individus hétérozygotes pour les deux loci (hétérozygotes doubles). 150 paires d'hétérozygotes doubles ont été placées dans la boîte à population dont les orifices de migration étaient ouverts. Au cours d'une journée, toute l'aire d'habitation de la boîte s'est trouvée peuplée, si bien que le nombre d'individus dans les compartiments variait de 1 à 37. Après quoi, on a placé dans les orifices de migration des tubes assurant un coefficient de migration ~0,03.

Deux éprouvettes contenant une nourriture de levure type se trouvaient en permanence dans chaque compartiment de la boîte à population. Au cours de toute la période de génération, les mouches d'une génération donnée s'envolaient de l'une des éprouvettes, dans la deuxième on voyait se développer les larves de la génération suivante. La fin de la période de génération était fixée en éloignant de la boîte tous les individus d'une génération donnée

avant l'envol des individus de la génération suivante. Après quoi, les éprouvettes d'où les mouches s'étaient envolées, étaient remplacées par des éprouvettes renfermant de la nourriture fraîche. La discontinuité des générations était ainsi assurée.

Afin de comparer les processus génétiques se déroulant dans les populations panmictique et subdivisée, on procédait parallèlement à une expérience dans une boîte à population témoin à un seul compartiment. Voici comment était formée la population panmictique témoin : avant l'envol des individus de la 1ère génération on plaçait une éprouvette supplémentaire avec de la nourriture fraîche dans chaque compartiment de la boîte divisée. A la fin de la 1ère génération, toutes les éprouvettes supplémentaires étaient retirées et utilisées pour former une population panmictique dans une boîte à un seul compartiment. On obtenait ainsi en pratique une identité parfaite de la composition génétique des populations panmictique et subdivisée. Dans l'expérience et le témoin, la nourriture était également la même.

Au cours de toute l'expérience ayant duré ~800 jours (>60 générations), les boîtes de population se trouvaient dans une chambre thermorégularisée où on maintenait une humidité et une température (24 ± 1 °C) constantes. La boîte divisée était éclairée régulièrement en dessus à travers un couvercle transparent de façon que les rayons lumineux tombent perpendiculairement à la direction de la migration. Grâce à ce procédé, malgré la phototaxie positive de *D. melanogaster*, une migration isotrope dans tout le système de subpopulations se trouvait assurée. Afin d'exclure une irrégularité non contrôlée de l'éclairage des différents compartiments, la boîte de population était placée sur une plaque tournante et au bout d'intervalles de temps déterminés on faisait tourner la boîte d'un angle arbitraire, grâce à quoi chaque compartiment occupait une position au hasard par rapport à l'espace environnant.

Dans les populations aussi bien subdivisée que panmictique on dénombrait à chaque génération l'effectif des individus et on enregistrait le rapport des sexes (sexratio). L'analyse électrophorétique des isoenzymes était effectuée toutes les 5 ou 10 générations. L'effectif de l'échantillon de chaque subpopulation était d'environ 40 individus, ce qui représente ~30 % du nombre total d'individus dans la subpopulation. L'échantillon tiré de la population panmictique était égal en moyenne à 100 individus. Au cours de l'expérience, ~9000 mouches ont été analysées.

PARTIE EXPÉRIMENTALE

Au fil des 50 premières générations, on déterminait, dans la population subdivisée et dans la panmixie, l'effectif des individus (fig. 3). L'effectif moyen de la population subdivisée était de 4036

individus (135 par subpopulation), et celui de la panmixie était de 4097 individus. Ainsi, dans des conditions de notre expérience, les effectifs des populations panmictique et subdivisée coïncidaient à peu près. Les variations de l'effectif dans les populations subdivisée et panmictique, dues à de légères fluctuations incontrôlables des conditions de l'environnement et, probablement, à la concurrence des larves pour la nourriture [23], sont aussi parfaitement de même phase. Ce fait reflète l'identité du patrimoine génétique collectif des deux populations et celle de l'influence de l'environnement sur

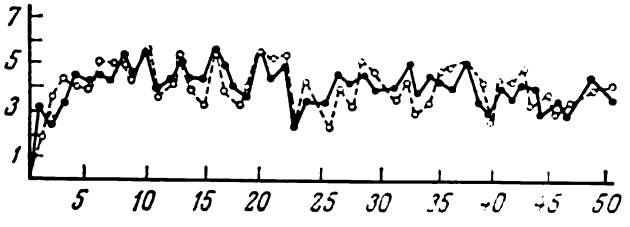


Fig. 3. Dynamique de l'effectif des mouches dans des générations successives non croisées de populations subdivisée (ligne continue) et panmictique (ligne pointillée) de *D. melanogaster*:
Sur l'axe des abscisses, les générations; sur l'axe des ordonnées, l'effectif en milliers d'individus

elles. Lors de l'estimation des effectifs, on a découvert non seulement des variations en temps correspondant à la figure 3, mais aussi des différences dans l'effectif des subpopulations de chaque génération donnée, c'est-à-dire qu'on observe une variabilité de l'effectif des composantes structurales du système tant dans le temps que dans l'espace.

L'enregistrement du rapport des sexes, fait tout au long de 30 générations, a montré qu'en cas de subdivision on observe chez quelques subpopulations des variations temporospatiales certaines du rapport des sexes avec une variance moyenne égale à 0,0035. Cependant, le sexratio moyen pris pour le système de subpopulations dans l'ensemble n'a pratiquement pas changé dans le temps et était de 0,5 ($\sigma^2 = 0{,}00025$) en accord parfait avec ce qui avait été décrit antérieurement pour les systèmes naturels de populations [3, 7].

L'identification électrophorétique des génotypes a été faite dans les 5e, 10e, 16e, 30e, 36e, 41e, 51e et 61e générations (*Tableau 1*). Les résultats typiques de l'analyse sont donnés sur la figure 4 où, sous forme de diagrammes sectoriels, on présente les distributions des fréquences des allèles F des loci Est-6 et α-PGDH pour les 16e, 36e et 61e générations. La forte dispersion des fréquences alléliques des deux loci témoigne d'une différenciation locale prononcée survenant déjà dans la 5e génération et qui devient très marquée à la 36e géné-

ration. Pour les 41e, 51e et 61e générations où les fréquences des allèles ne variaient pratiquement pas, on a comparé les distributions globales des génotypes calculées en vertu de l'équation de Hardy-Weinberg. La comparaison effectuée a montré que ce qui est typique du système dans son ensemble, c'est un déficit significatif des hétérozygotes comme conséquence de l'effet Walund *.

Tableau 1

Principaux paramètres statistiques des variances de fréquences des allèles des loci α-PGDH et Est-6 dans les générations successives d'un système expérimental de subpopulations de *D. melanogaster*

Locus	Génération	k	N	N'	$\bar{p}$	σ^2_p	Test à l'homogénéité de la population, $\chi^2_{0,001}$
α-PGDH	5	27	4203	1109	0,36	0,0122	$616,19 > \chi^2(26) = 54,1$
	10	23	5315	943	0,41	0,0340	$1250,17 > \chi^2(22) = 48,3$
	16	29	5447	1215	0,40	0,0485	$2000,03 > \chi^2(28) = 56,9$
	30	29	4163	1227	0,51	0,0819	$2629,00 > \chi^2(28) = 56,9$
	36	30	4525	1198	0,57	0,0726	$2683,00 > \chi^2(29) = 58,3$
	41	29	3835	1149	0,62	0,0491	$1550,99 > \chi^2(28) = 56,9$
	51	30	3571	1097	0,62	0,0549	$1659,84 > \chi^2(29) = 58,3$
	61	30	3605	949	0,65	0,0717	$1650,39 > \chi^2(29) = 58,3$
Est-6	5	28	4203	1196	0,50	0,0121	$569,61 > \chi^2(27) = 55,5$
	10	29	5315	1367	0,42	0,0146	$629,53 > \chi^2(28) = 56,9$
	16	28	5447	1149	0,46	0,0119	$488,48 > \chi^2(27) = 55,5$
	30	27	4163	1228	0,48	0,0453	$1353,11 > \chi^2(26) = 54,1$
	36	28	4525	1166	0,44	0,0409	$1373,85 > \chi^2(27) = 55,5$
	41	29	3835	1212	0,42	0,0440	$1360,42 > \chi^2(28) = 56,9$
	51	25	3571	757	0,34	0,0299	$739,85 > \chi^2(24) = 51,2$
	61	28	3605	853	0,38	0,0610	$1273,99 > \chi^2(27) = 55,5$

Remarque. k est le nombre de subpopulations étudiées,
N l'effectif des individus dans le système,
N' l'effectif des individus étudiés,
$\bar{p}$ la fréquence moyenne de l'allèle F,
σ^2_p la variance des fréquences d'allèles entre populations.

Au *Tableau 1* sont mentionnés les principaux paramètres, admis en génétique des populations, d'une population subdivisée, notés au cours de toute l'expérience et permettant une estimation quantitative de la variabilité observée. Nous voyons que les fréquences moyennes des allèles du locus Est-6 ne manifestent pas de variations dirigées

* Dans les générations indiquées, les valeurs de $\chi^2 \geqslant 53,26$ pour l'α-PGDH et de $\chi^2 \geqslant 18,24$ pour l'Est-6, le nombre de degrés de liberté étant égal à 1.

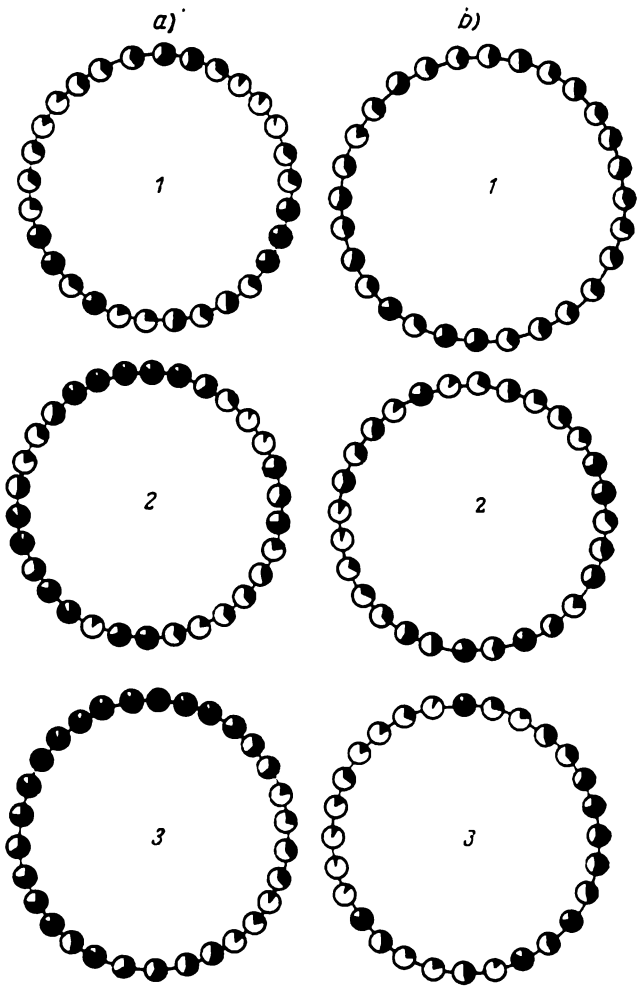


Fig. 4. Différenciation locale des fréquences géniques pour les loci α-PGDH (*a*) et Est-6 (*b*) dans les générations successives d'une population subdivisée de *D. melanogaster*:

1-3 — respectivement 36e, 51e et 61e générations.
Les parties obscurcies des cercles correspondent aux valeurs des fréquences des allèles rapides F dans chacune des subpopulations

dans le temps, mais oscillent autour d'une valeur égale à 0,43. La fréquence moyenne de l'allèle F du locus α-PGDH dans l'intervalle entre la 5e et la 41e génération augmente, ce qui atteste la présence dans notre système d'une sélection dirigée pour un locus donné. Les questions rattachées à l'effet de la sélection dans le système de

subpopulations seront l'objet d'une communication particulière. Ici, il faut accorder une grande attention à l'hétérogénéité fortement prononcée du patrimoine génétique collectif de la population subdivisée pour les deux loci étudiés, ce qui reflète apparemment une forte influence de la dérive génétique. Etant donné que la présence de processus stochastiques est en rapport avec le petit effectif de la population, une manifestation si nette de la dérive génétique dans notre système où l'effectif moyen de la subpopulation est assez élevé ($\bar{N} = 135$ individus), peut sembler, à première vue, inattendue. Cependant, il faut tenir compte de ce que la grandeur reproductive effective (N_e) de la population est, généralement, de beaucoup inférieure à son effectif réel. Compte tenu des données connues et aussi

Tableau 2

Valeurs théoriques et expérimentales des variances des fréquences géniques dans un système de populations de *D. melanogaster*

Locus	Génération	$\bar{p}$	r (1)	$\sigma^2_{p\ (théor)}$	$\sigma^2_{p\ (obs)}$
α-PGDH	51	0,62	0,72	0,0879	0,0549
	61	0,65	0,81	0,1063	0,0717
Est-6	51	0,34	0,31	0,0437	0,0299
	61	0,38	0,24	0,0424	0,0610

Remarque. $\bar{p}$ est la fréquence moyenne de l'allèle F,
r (1) la corrélation des fréquences géniques entre les subpopulations connexes,
$\sigma^2_{p\ (théor)}$ la valeur théorique de la variance,
$\sigma^2_{p\ (obs)}$ la valeur observée de la variance.
Les calculs étaient faits pour $N_e = 50$ et un coefficient de migration $m = 0{,}03$.

de certaines particularités écologiques propres à notre système *, la grandeur reproductive effective de la subpopulation doit être au moins de deux fois inférieure à la valeur que nous avons calculée comme moyenne harmonique d'une série de populations (120 individus). Ceci est confirmé par l'analyse comparée des variances théoriques et observées des fréquences d'allèles. Le *Tableau 2* montre les

* Dans l'expérience de Buri [12], effectuée sur de petites populations de *D. melanogaster*, N_e était respectivement 56 et 72 % de l'effectif dans les séries à grande et moindre densité de la population, ce que Buri rattache à une densité élevée des larves dans la nourriture. Dans ces conditions, l'avantage dans la production de la descendance revient aux femelles qui ont été les premières à pondre des œufs [26]. Remarquons que dans le travail de Buri [12], toutes les femelles étaient alimentées en même temps. Dans notre expérience où la ponte des œufs a lieu durant tout l'envol, la différence entre N_e et l'effectif réel doit augmenter, étant donné que tous les individus n'ont pas le temps de laisser une descendance.

valeurs expérimentales et théoriques des variances obtenues pour la 51e et la 61e génération en admettant que $N_e = 50$. Les calculs étaient effectués suivant la formule déduite par Kimura et Weiss [18] pour une population en équilibre sans sélection :

$$\sigma_{\bar{p}}^2 = \frac{\bar{p}(1-\bar{p})}{1+4N_e m\,[1-r\,(1)]}$$

où $\sigma_{\bar{p}}^2$ est la variance entre populations,
p la fréquence génique moyenne,
m le coefficient de migration,
r (1) la corrélation des fréquences géniques entre les subpopulations connexes.

Comme le montre le *Tableau 2*, les valeurs expérimentales et théoriques des variances pour $N_e = 50$ sont du même ordre de grandeur, bien que dans 3 cas sur 4 les valeurs expérimentales soient tout de même inférieures aux valeurs théoriques. Ceci est probablement dû à l'effet de la sélection.

L'apparition d'une différenciation locale dans la population subdivisée dépend non pas tellement de la valeur de N_e que de son produit par le coefficient de migration représentant le nombre de migrants venus dans la subpopulation au cours d'une génération. La différenciation locale des fréquences géniques dans le système de subpopulations ne survient que si un paramètre donné est d'une valeur négligeable. La condition d'apparition d'une différenciation locale pour un modèle échelonné fermé unidimensionnel [21] est

$$N_e m < k/\pi^2$$

où m est le coefficient de migration,
k le nombre de subpopulations.

Dans notre système $m = 0{,}03$, $k = 30$. Pour $N_e \cong 50$, la condition $N_e m < 3{,}3$ est remplie.

DISCUSSION

Déjà le fait que, dans la Nature, les populations réelles occupent de vastes aires, permet de supposer en elles l'absence de panmixie, étant donné que l'augmentation des distances entre individus réduit la probabilité de leur rencontre et de leur croisement [1-3, 5]. L'existence de barrières géographiques et comportementales provoque l'apparition de systèmes de subpopulations en partie isolées [10] et, comme conséquence, l'hétérogénéité génétique spatiale. En effet, de nombreux chercheurs [25, 27, 29] signalent l'existence d'une différenciation locale des fréquences géniques dans les populations naturelles.

Il y a deux causes de la variabilité géographique des fréquences géniques [13]: 1° la présence dans l'aire de plusieurs niches écologiques dans lesquelles s'exerce une sélection de direction ou d'intensité différentes (sélection en milieu hétérogène); 2° l'interaction de la dérive et de la migration dans le système de subpopulations. Au cours de l'analyse de la variabilité génétique des populations naturelles, il n'est pas facile de choisir entre ces deux situations, car, dans la Nature, il est pratiquement impossible d'éliminer le facteur sélection en raison de la grande diversité des conditions de l'environnement [19]. Les recherches théoriques ne fournissent pas non plus de réponse précise à cette question. Certains auteurs affirment [17] qu'en présence de valeurs suffisamment faibles du coefficient de migration, le caractère subdivisé des populations exerce une influence sur les processus génétiques qui s'y déroulent, ce qui suscite une différenciation locale des fréquences géniques. Selon d'autres [22], même une migration minimale (1 migrant par génération) conduit à l'identité de la génétique des populations subdivisées et panmictiques. De plus, les résultats obtenus sur des modèles mathématiques élémentaires sont souvent généralisés aux plus compliqués et même à n'importe quelle population subdivisée. Quelquefois, en se basant sur des calculs analytiques, on tire catégoriquement la conclusion que la cause de la différenciation locale des fréquences géniques est toujours la pression de sélection en milieu hétérogène [28].

Comme le montrent les particularités méthodiques de l'expérience que nous avons entreprise, les conditions de l'environnement sont parfaitement identiques pour toutes les subpopulations, c'est-à-dire que la possibilité d'une sélection en milieu hétérogène est tout à fait exclue. Bien plus, comme le montrent les publications [11, 30], les deux loci étudiés par nous sont soumis à la pression de formes de sélection telles (dirigée pour l'α-PGDH et dépendante de la fréquence ou stabilisante pour l'Est-6) qu'elles doivent égaliser les fréquences géniques dans les subpopulations. Cependant, dans notre expérience, nous observons un niveau élevé de différenciation locale, ce qu'attestent, en particulier, les valeurs des variances entre populations (à la 61e génération pour l'Est-6 $\sigma^2 = 0{,}0610$, pour l'α-PGDH $\sigma^2 = 0{,}0717$).

Ces données s'accordent bien avec les résultats des études de populations naturelles de certaines espèces animales [3, 6, 9] et de l'homme [24] lorsque les chercheurs ne se bornaient pas à l'analyse d'échantillons pris au hasard, mais pouvaient obtenir des preuves suffisamment fiables de l'influence du facteur isolement sur la distribution temporospatiale des fréquences géniques. Dans le travail d'Y. Altoukhov et de ses collaborateurs [9] ayant étudié la variabilité génétique dans un système de saumons rouges d'après les loci codant pour la phosphoglucomutase et la lactate-déshydrogénase, la variance entre populations était d'un ordre de grandeur plus petite

que dans notre expérience (0,0029 pour PGM et 0,0067 pour LDH). Cependant, elle était tout à fait fiable. Dans l'étude du polymorphisme du dessin de la coquille du mollusque terrestre *Chondrus bidens* Kryn. [6], l'estimation quantitative de la différenciation donne une valeur de la variance entre populations à peu près du même ordre de grandeur que dans notre expérience. Une hétérogénéité prononcée des fréquences géniques ($\sigma^2 = 0{,}008 - 0{,}023$) a également été découverte en étudiant la distribution des groupes sanguins des systèmes MN et ABO dans les populations aborigènes de Sibérie [24]. Signalons que les loci envisagés dans ces travaux ne sont pas neutres. Cependant, la différenciation locale des fréquences géniques dans les systèmes de populations se maintient fermement malgré l'effet de la sélection.

Les données que nous avons obtenues sur la population subdivisée expérimentale de *D. melanogaster* de même que celles mentionnées plus haut sur les populations naturelles d'espèces diverses permettent de conclure que la différenciation spatiale des fréquences géniques dans le système de subpopulations apparaît et se maintient même dans les conditions d'une pression de sélection dirigée ou stabilisante.

RÉFÉRENCES

[1] *Altukhov Yu. P.* Recent physiological, biochemical and immunological studies on the problem of intraspecific differentiation in marine fish. Special Meet. Biochem. Serol. Identification, Fish Stocks, Dublin, 1969. *Rapp. Proc. Verb.*, 1971, 161, 103.

[2] *Altukhov Yu. P.* Rational management of marine biological resources in the light of population genetics. In : Propog. marine resources Pacific Ocean, Tokyo, Tokai Univ., 1973, 31-39.

[3] *Алтухов Ю. П.* Популяционная генетика рыб. М., Пищевая пром-сть, 1974.

[4] *Алтухов Ю. П., Бернашевская А. Г.* Экспериментальное моделирование динамики генных частот в системе полуизолированных популяций. *Докл. АН СССР*, 1978, 238, 3, 712.

[5] *Алтухов Ю. П., Лиманский В. В., Паюсова А. Н., Трувеллер К. А.* Иммунологический анализ внутривидовой дифференцировки европейского анчоуса, обитающего в Черном и Азовском морях. Сообщение II. Элементарные популяции анчоуса и их место в генетико-популяционной структуре вида. *Генетика*, 1969, 5, 5, 81.

[6] *Алтухов Ю. П., Лившиц Г. М.* Факторы дифференциации и интеграции генофонда изолированной популяции моллюска *Chondrus bidens* (Kryn.). *Докл. АН СССР*, 1978, 238, 4, 44.

[7] *Алтухов Ю. П., Победоносцева Е. Ю.* Экспериментальное моделирование генетических процессов в подразделенных популяциях. *Докл. АН СССР*, 1978, 238, 2, 466.

[8] *Алтухов Ю. П., Салменкова Е. А., Омельченко В. Т., Сачко Г. Д., Самнько В. И.* О числе мономорфных и полиморфных локусов в популяции кеты *Oncorhynchus keta* Walb. — одного из тетраплоидных видов лососевых. *Генетика*, 1972, 8, 2, 67.

[9] *Алтухов Ю. П., Салменкова Е. А., Коновалов С. М., Пудовкин А. И.* Стационарность распределений частот генов лактатдегидрогеназы и

фосфоглюкомутазы в системе субпопуляций локального стада рыб (на примере *Oncorhynchus nerka* Walb.). Сообщение I. Стабильность стада в поколениях при одновременной изменчивости слагающих их структуру субпопуляций. *Генетика*, 1975, 11, 4, 44.

[10] *Алтухов Ю. П., Рычков Ю. Г.* Популяционные системы и их структурные компоненты. Генетическая стабильность и изменчивость. *Ж. общ. Биол.*, 1970, 31, 5, 507.

[11] *Barger E. M.* A temporal survey of allelic variation in natural and laboratory populations of *Drosophila melanogaster*. *Genetics*, 1971, 67, 121.

[12] *Buri P.* Gene frequency in small populations of mutant *Drosophila*. *Evolution*, 1956, 10, 367.

[13] *Cavalli-Sforza L. L., Bodmer W. F.* The genetics of human populations. San-Francisco, Freeman, 1971.

[14] *Dobzhansky Th.* Genetics of the evolutionary process. N.Y.-L., Columbia Univ. Press, 1970.

[15] *Дубинин Н. П.* Эволюция популяций и радиация. М., Атомиздат, 1966.

[16] *Kimura M.* Genetic variability in a finite population due to mutational production of neutral isoalleles. *Genet. Res.*, 1968, 11, 247.

[17] *Kimura M., Maruyama T.* Pattern of neutral polymorphism in a geographically structured population. *Genet. Res.*, 1971, 18, 125.

[18] *Kimura M., Weiss G. H.* The stepping-stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, 1964, 49, 561.

[19] *Lewontin R. C.* The genetic basis of evolutionary change. N.Y.-L., Columbia Univ. Press, 1974.

[20] *Li C. C.* First course in population genetics. California, Boxwood Press, Pacific Grove, 1976.

[21] *Maruyama T.* On the rate of decrease of heterozygosity in circular stepping-stone models of populations. *Theor. Populat. Biol.*, 1970, 1, 101.

[22] *Moran P. A. P.* The theory of some genetical effects of population subdivision. *Austral. J. Biol. Sci.*, 1959, 12, 109.

[23] *Nogués R. M.* Population size fluctuations in the evolution of experimental cultures of *Drosophila subobscura*. *Evolution*, 1977, 71, 200.

[24] *Рычков Ю. Г., Русакова Л. Л., Раппопорт М. Н., Шереметьева В. А.* Факторы генетической дифференциации популяционной системы коренного населения Северной Азии. Сообщение I. Определение интегрирующих значений адаптивной ценности генотипов групп крови АВО в коренном населении Сибири. *Генетика*, 1973, 9, 2, 136.

[25] *Ryman N., Beckman G., Brunn-Petersen G., Reuterwall C.* Variability of red cell enzymes and genetic implications of management policies in Scandinavian moose (*Alces alces*). *Hereditas*, 1977, 85, 2, 157.

[26] *Sang J. H.* The ecological determinants of population growth in a *Drosophila* culture. IV. The significance of successive batches of larvae. *Physiol. Zool.*, 1949, 22, 202.

[27] *Selander R. K., Kaufman D. W.* Genetic structure of population of the brown snail (*Helix aspersa*). *Evolution*, 1975, 29, 385.

[28] *Spleth P. T.* Gene flow and population differentiation. *Genetics*, 1974, 78, 968.

[29] *Tabachnick W. J.* Geographic variation of five biochemical polymorphisms in *Notophthalmus viridescens*. *J. Heredity*, 1977, 68, 2, 117.

[30] *Yarbrough K. M., Kojima K.* The mode of selection at the polymorphic esterasa-6 locus in cage populations of *D. melanogaster*. *Genetics*, 1967, 47, 677.

[31] *Wright S.* The genetical structure of populations. *Ann. Eugenics*, 1951, 15, 323.

SIMULATION EXPÉRIMENTALE DES PROCESSUS GÉNÉTIQUES DANS UN SYSTÈME DE POPULATIONS DE *DROSOPHILA MELANOGASTER* CORRESPONDANT AU MODÈLE ÉCHELONNÉ CIRCULAIRE

Communication II. Stabilité de la composition des allèles et dépendance périodique de la variabilité des fréquences d'allèles par rapport à la distance

Y. Altoukhov et A. Bernachevskaïa

Dans la communication précédente, nous avons justifié l'approche de la simulation expérimentale des phénomènes génétiques dans un système de populations partiellement isolées de *D. melanogaster*, organisé selon le type du modèle échelonné, et nous avons montré la différenciation locale stable des fréquences d'allèles des loci α-PGDH et Est-6. Dans la présente communication, nous examinerons la dynamique génétique dans le temps de cette population subdivisée et nous montrerons que dans un système de ce genre la différenciation locale n'est pas fortuite, mais qu'elle se rapproche d'une différenciation périodique reflétant, peut-être, un processus stationnaire de réorganisation de la structure génétique de la population dans les générations.

MATÉRIEL ET MÉTHODE

Nous avons utilisé dans l'expérience une population subdivisée de *D. melanogaster* organisée selon le type du modèle échelonné circulaire (stepping-stone) [4]. Les subpopulations réunies par des canaux de migration formaient un anneau, l'échange d'individus ne se réalisant qu'entre les subpopulations connexes. Le système se composait de 30 subpopulations avec un coefficient moyen de migration égal à $0{,}027 \pm 0{,}007$. L'effectif moyen de chaque subpopulation était égal à 135. Cependant, la grandeur reproductive effective dans les conditions de l'expérience était beaucoup plus basse, de l'ordre de 50 individus [1]. Les générations ne se superposaient pas.

La boîte à population se trouvait dans une chambre thermorégularisée où on maintenait une humidité et une température (24 ± 1 °C) constantes et un éclairage régulier. Au cours de l'expérience une attention particulière a été apportée à assurer l'absence de sélection dans le milieu hétérogène. Dans ce but, on a créé pour toutes les subpopulations des conditions identiques du milieu.

© Издательство Наука, 1981.

En qualité de marqueurs génétiques on a utilisé des loci codominants autosomiques, codant pour l'α-PGDH et l'Est-6. La population a été fondée avec une fréquence initiale des allèles rapide (F) et lent (S) des deux loci égale à 0,5. Toutes les 5 ou 10 générations à partir de la 5e, on effectuait une analyse électrophorétique pour identifier les génotypes et déterminer les fréquences géniques. L'échantillon tiré de chaque subpopulation était égal en moyenne à 40 individus. Une expérience parallèle se faisait avec une population panmictique témoin dont l'effectif était le même que dans la population d'expérience. Les conditions de formation de la population panmictique assuraient également l'identité de son patrimoine génétique à celui de la population subdivisée. Tous les détails de notre méthode sont exposés dans la communication précédente.

PARTIE EXPÉRIMENTALE

Un des importants résultats obtenus dans le présent travail (*Tableau 1*) est l'obtention par le système de populations expérimental

Tableau 1

Dynamique des fréquences des allèles des loci α-PGDH et Est-6 dans une population subdivisée expérimentale de *Drosophila melanogaster*

Locus	Génération	k	N	$\bar{p}$	$\sigma^2_{\bar{p}}$	Test à l'homogénéite des dispersions, $\chi^2_{0,001}$ (avec degrés de liberté)
α-PGDH	5	27	41	0,36	0,0122	Générations 5e-61e
	10	23	41	0,41	0,0340	$26{,}00 > \chi^2(7) = 24{,}3$
	16	29	42	0,40	0,0485	
	30	29	42	0,51	0,0819	
	36	30	40	0,57	0,0726	Générations 36e-61e
	41	29	40	0,62	0,0491	$2{,}48 < \chi^2(4) = 18{,}8$
	51	30	37	0,62	0,0549	
	61	30	32	0,65	0,0717	
Est-6	5	28	43	0,50	0,0121	Générations 5e-61e
	10	29	47	0,42	0,0146	$36{,}77 > \chi^2(7) = 24{,}3$
	16	28	41	0,46	0,0119	
	30	27	45	0,48	0,0453	
	36	28	42	0,44	0,0409	Générations 36e-61e
	41	29	42	0,42	0,0440	$3{,}22 < \chi^2(4) = 18{,}8$
	51	25	30	0,34	0,0299	
	61	28	30	0,38	0,0610	

Remarque. k est le nombre de subpopulations étudiées,
N l'effectif moyen d'un échantillon tiré d'une subpopulation,
$\bar{p}$ la fréquence moyenne de l'allèle F,
$\sigma^2_{\bar{p}}$ la variance des fréquences d'allèles entre populations.

d'un état stable par rapport à la fréquence génique, tant pour l'Est-6 que pour l'α-PGDH. De plus, la situation concernant l'α-PGDH est particulièrement intéressante. En effet, si l'on suit la dynamique des fréquences d'allèles du locus Est-6 pour l'expérience et le témoin (fig. 1, *b*), il apparaît qu'il n'existe pas de différences significatives entre ces populations: dans l'un et l'autre cas, une fréquence proche de 0,45 est entretenue, ce qui reflète la pression de sélection stabilisante ou de sélection proche de cette dernière. Il est évident que dans ces conditions on ne puisse rien dire de l'importance de la structure de populations.

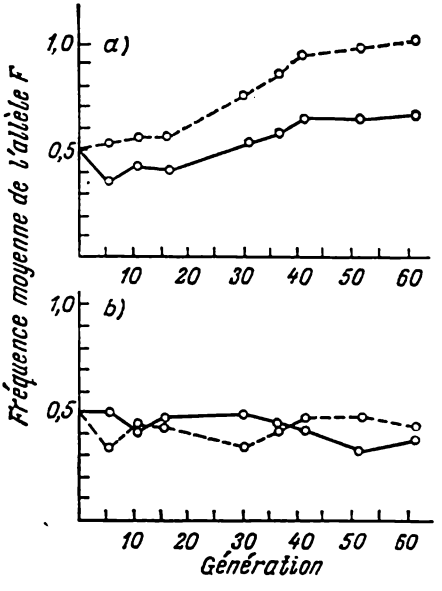


Fig. 1. Dynamique des fréquences d'allèles des loci α-PGDH (*a*) et Est-6 (*b*) dans les générations successives: La ligne continue et la ligne pointillée représentent les fréquences moyennes de l'allèle F respectivement dans des populations subdivisée et panmictique

Au contraire, la dynamique des fréquences d'allèles du locus α-PGDH (fig. 1, *a*) est déterminée par la pression de sélection dirigée et nous voyons que la population panmictique témoin évolue rapidement parvenant, au bout de 60 générations, à une fréquence de l'allèle F égale à 0,98, ce qui correspond à $\Delta q = 0{,}008$ par génération. Dans la population subdivisée expérimentale on observe également un écart vectorisé de la concentration de l'allèle F. Cependant, la cadence de ce processus est fortement ralentie en comparaison du témoin et ne représente que 0,0025 par génération. Ainsi, sous la pression de sélection dirigée, la cadence de l'évolution de la population subdivisée se trouve au moins de trois fois inférieure à celle de la panmixie.

On comprend que, même dans un ensemble arbitraire de populations, seule la dérive génétique est capable de ralentir son évolution dirigée, si l'on considère cet ensemble comme un tout. Un tel effet avait déjà été démontré par nous dans d'autres travaux de ce cycle [2, 3]. Mais en l'occurrence, nous avons affaire non pas à un ensemble arbitraire de petits isolats, mais à un système de subpopulations constamment en interaction les unes avec les autres par échange de gènes. Il est possible que justement cette circonstance explique le fait qu'à partir de la 41e génération et jusqu'à la fin de l'expérience,

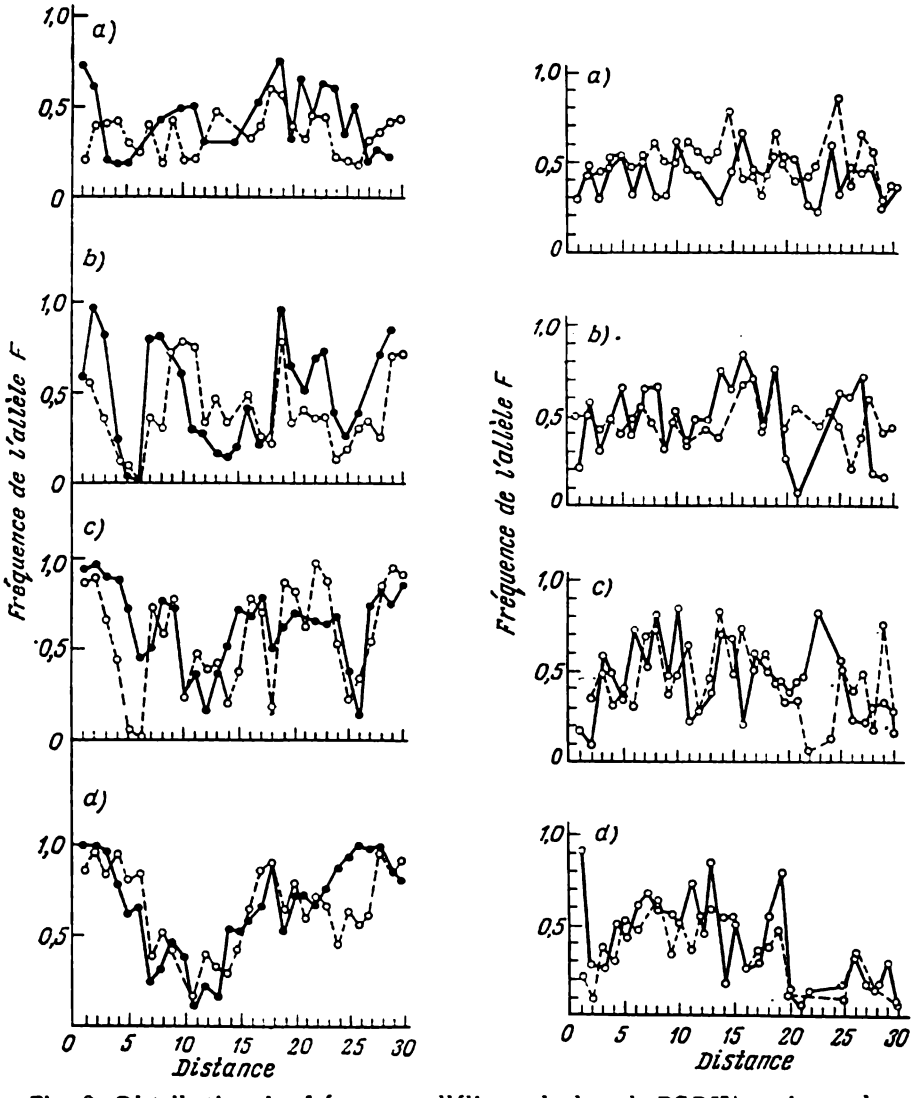


Fig. 2. Distribution des fréquences alléliques du locus α-PGDH au niveau des subpopulations dans les générations successives d'une population subdivisée expérimentale de *D. melanogaster*:

a, *b*, *c*, *d* — respectivement 5e-10e, 15e-30e, 36e-41e, 51e-61e générations

Fig. 3. Distribution des fréquences alléliques du locus Est-6 au niveau des subpopulations dans les générations successives d'une population subdivisée expérimentale de *D. melanogaster*:

Les désignations sont les mêmes que sur la figure 2

c'est-à-dire au fil de 20 générations, la fréquence moyenne de l'allèle F de l'α-PGDH et la variance entre groupes des fréquences dans la population subdivisée restent pratiquement inchangées (v. *Tableau 1* et figure 1, *b*).

En effet, si nous examinions la variabilité génétique au niveau des subpopulations dans une série de « tranches » de temps consécutives, nous verrions une tendance nette à la formation par le système d'une structure interne stable (fig. 2 et 3). Ces figures montrent les dépendances des fréquences des allèles F de l'α-PGDH et de l'Est-6 par rapport à la distance selon l'axe de migration pour une série de générations. L'axe de migration est une ligne réunissant des subpopulations disposées à la suite les unes des autres. Les chiffres sur l'axe de migration représentent la distance en pas, chaque pas correspondant au passage de la subpopulation précédente à la suivante. Pour faciliter l'analyse sur les figures 2 et 3, les générations sont présentées par deux: la ligne pointillée et la ligne continue correspondent respectivement à la génération précédente et à la suivante. La fréquence initiale de l'allèle F pour les deux loci était donnée égale à 0,5 pour toutes les subpopulations.

Examinons d'abord la dynamique des fréquences géniques dans les subpopulations pour le locus α-PGDH (v. fig. 2). Dans les premières générations de l'expérience (v. fig. 2, *a*), la variabilité observée est d'un caractère aléatoire. Cependant, dès la 10e génération une tendance typique commence à se manifester: sur les courbes apparaissent des pointes et des creux réunis par des zones de fréquences intermédiaires, c'est-à-dire qu'une certaine périodicité s'ébauche dans la dépendance des fréquences d'allèles par rapport à la distance entre subpopulations. Au cours de l'expérience, le caractère périodique de la variabilité génétique spatiale pour le locus α-PGDH devient encore plus net: la disposition des points figurant les fréquences dans certaines subpopulations devient plus rythmique.

La périodicité de la variation des fréquences géniques pour le locus α-PGDH est particulièrement nette dans l'illustration comparant les dépendances obtenues expérimentalement et la fonction périodique élémentaire $pF = a + b\,(\cos l/\lambda + 1)$ où a et b sont des coefficients numériques, l est la distance en pas sur l'axe de migration, λ la période de variation des fréquences le long de l'axe de migration.

L'approximation des valeurs expérimentales des fréquences géniques par la fonction indiquée a donné les résultats suivants:

$$pF_{36} = 0{,}1 + 0{,}40\left(\cos 2\pi \frac{l}{10} + 1\right),$$

$$pF_{41} = 0{,}1 + 0{,}42\left(\sin 2\pi \frac{l-3{,}5}{15} + 1\right),$$

$$pF_{51} = 0{,}1 + 0{,}45\left(\cos 2\pi \frac{l}{30} + 1\right).$$

On voit que la valeur de la période de variation des fréquences d'allèles suivant l'axe de migration pour les 36e, 41e et 51e générations est dans le rapport de 1 : 2 : 3. De la 51e à la 61e génération, la période n'a pas changé (fig. 4, *a*, *b* et *c*). La comparaison des données présentées sur la figure 4, *b* et *c* montre qu'à la 61e génération le passage de deux périodes à une seule n'était pas réalisé complètement. La pointe de la courbe expérimentale à la figure 4, *b* correspondant aux fréquences dans les 15e-21e subpopulations ne s'est aplatie que partiellement, et les valeurs des fréquences géniques sur ce secteur de l'aire sont disposées au-dessus de la courbe théorique approximative.

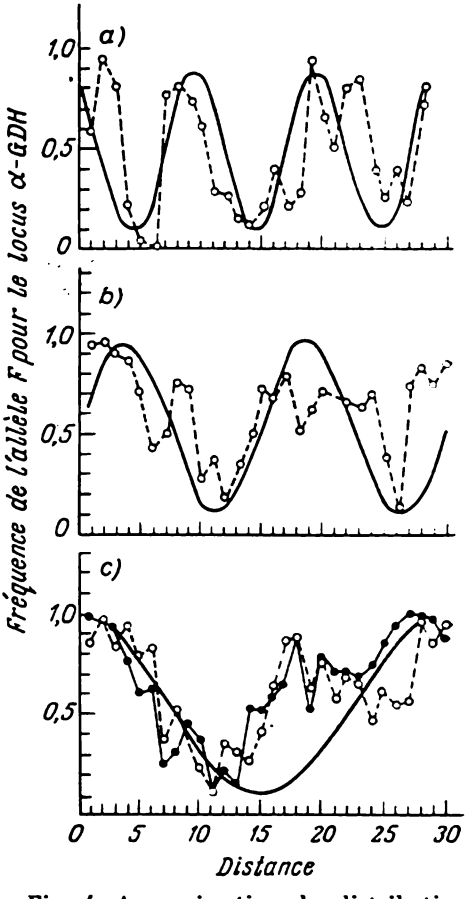


Fig. 4. Approximation des distributions expérimentales des fréquences alléliques du locus α-PGDH par une fonction périodique :
a, b, c — 36e-41e et 51e-61e générations. (V. explications dans le texte)

Ainsi, en se basant sur le caractère de la variation de la différenciation spatiale des fréquences géniques pour le locus α-PGDH, l'« histoire » du système de populations étudié par nous peut être conventionnellement divisée en trois étapes : 1° formation de la structure périodique (1e-16e générations) ; 2° augmentation de la période de variation des fréquences suivant l'axe de migration (30e-51e générations) ; 3° stade stationnaire ou quasi stationnaire (51e-61e générations). On peut en tirer la conclusion préalable que la présence de plusieurs maxima de fréquences le long de l'axe de migration typique des stades initiaux de développement de la structure du système est un état instable et que le système cherche à établir une seule période de variations des fréquences.

Passons maintenant à l'examen de la variabilité spatiale des fréquences géniques du second locus étudié par nous : Est-6 (v. fig. 3). Dans les premières générations de l'expérience, la variabilité génétique spatiale pour ce locus, de même que dans le cas du locus α-PGDH, est purement fortuite. Dans la suite, son caractère ne se modifie pas sensiblement. C'est seulement à la fin de l'expérience (à partir de la 50e génération) qu'apparaissent sur la courbe, caractérisant la dépendance entre les variations de fréquences géniques et la distance, un creux assez stable (20e-30e subpopulations) et une pointe (5e-15e subpopulations) qui se maintiennent dans un intervalle de 10 générations (v. fig. 3, *d*).

Ainsi, la variabilité génétique spatiale pour le locus Est-6 marque également une tendance à la formation d'une certaine structure non fortuite stable, bien que moins prononcée que dans le cas d'α-PGDH.

Ces conjectures qualitatives sur la tendance du système de subpopulations à atteindre un état stable se trouvent confirmées par l'analyse corrélative des données obtenues. L'estimation que nous avons faite des corrélations des fréquences géniques entre les paires de générations (*Tableau 2*) a révélé une croissance des valeurs des corré-

Tableau 2

Corrélations des fréquences géniques au niveau des subpopulations entre les paires de générations de la population subdivisée

Locus	Génération				
	5e et 16^{e}	16e et 30^{e}	30e et 41^{e}	41e et 51^{e}	51e et 61^{e}
α-PGDH	0,10	0,48	0,38	0,64	0,77
Est-6	0,07	0,21	0,16	0,53	0,52

lations dans le temps pour les deux loci, ce qui atteste l'augmentation de la stabilité de la structure du système.

L'étude théorique du modèle échelonné ouvert unidimensionnel, effectuée par Kimura et Weiss [5] d'une façon analytique, a montré que dans le cas d'un état équilibré du système la corrélation entre les fréquences géniques des subpopulations dans une génération donnée décroît en fonction exponentielle avec la distance. Cette dépendance est l'expression du fait évident qu'à mesure de l'augmentation de la distance entre les subpopulations, la possibilité d'un échange de gènes entre elles diminue fortement dans un modèle échelonné unidimensionnel. Nous avons étudié la dépendance de la

corrélation des fréquences de l'allèle F du locus α-PGDH par rapport à la distance entre les subpopulations dans la 61e génération. Les résultats de l'analyse sont présentés graphiquement sur la figure 5. Au lieu d'une décroissance exponentielle monotone, dans les conditions de notre expérience on observe une décroissance initiale de la corrélation jusqu'à une valeur minimale au milieu de l'axe de migration suivie d'une montée symétrique. Dans ce cas, le caractère symétrique de la variation de corrélation est dû à ce que le système étudié par nous est fermé. En nous déplaçant suivant la courbe fermée jusqu'au milieu de l'axe de migration, nous nous éloignons de la subpopulation que nous avons choisie comme point de référence, et la probabilité d'un échange de gènes avec elle va en diminuant Cependant, à partir du milieu de l'axe, nous nous rapprochons de la subpopulation initiale du côté opposé et la probabilité d'un échange de gènes commence à croître selon une loi analogue.

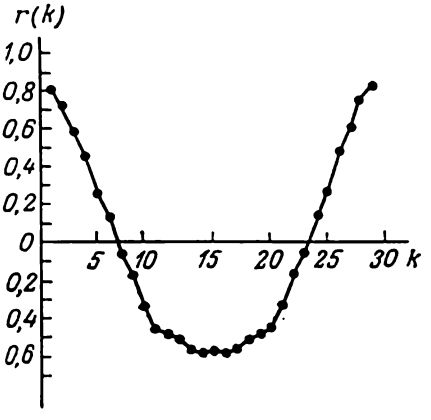


Fig. 5. Dépendance de la corrélation des fréquences alléliques du locus α-PGDH par rapport à la distance dans la 61e génération de l'expérience

DISCUSSION

L'expérience que nous avons effectuée montre que la particularité typique de la dynamique des fréquences géniques dans un système de subpopulations, organisé selon le type d'un modèle échelonné fermé unidimensionnel, est la périodicité. Les variations périodiques de fréquences géniques pour différents loci suivant la distance géographique avaient été notées par Rytchkov et Chérémétiéva [7] en étudiant les populations aborigènes circumpolaires de l'Eurasie. Le système étudié par Rytchkov correspond à peu près à un modèle échelonné non fermé (linéaire) unidimensionnel ayant un coefficient de migration ~0,03 et un effectif de subpopulation égal à 45, c'est-à-dire que les paramètres de la structure sont très proches de ceux de notre population expérimentale.

Transportée sur un plan, la dépendance périodique de la fréquence génique par rapport à la distance que nous avons notée plus haut correspond à la surface en relief des fréquences avec alternance des pointes et des creux, c'est-à-dire des domaines de maxi-

mums et de minimums. Un tableau analogue de la variabilité a été décrit par R. Selander [8] pour une population de souris de maison (*Mus musculus*) où la distribution spatiale des fréquences alléliques des loci, codant pour les enzymes, manifestait un caractère mosaïque avec passage rythmique entre les domaines de maximums et de minimums. L'effectif de la colonie (tribu) de souris de maison est égal en moyenne à 20 individus. En raison de leurs particularités comportementales, la migration entre les tribus est très faible. R. Selander interprète l'apparition du type décrit de différenciation locale des fréquences géniques comme le résultat de la subdivision de la population, étant donné que l'environnement dans le grenier où se trouvaient les colonies de souris, était parfaitement homogène.

Une variabilité génétique spatiale à caractère mosaïque a été également obtenue par simulation sur machine électronique d'une population dont la structure de migration correspondait à l'isolement dit par distance [6]. Il est intéressant de noter que, lors de la simulation sur ordinateur, la structure de populations obtenue était stable dans le temps: les domaines de pointes et de creux des fréquences se maintenaient durant des dizaines de générations. Un tel maintien des domaines de maximums et de minimums des fréquences géniques s'observait également dans notre expérience. Apparemment, la stabilité de la structure interne de populations pendant de nombreuses générations reflète un processus stationnaire de réorganisation de la structure génétique dans un système de subpopulations isolées en partie.

En même temps, on observe une différence dans le caractère de la variabilité spatiale des fréquences alléliques des loci α-PGDH et Est-6 se traduisant par une structure spatiale déterminée pour le locus α-PGDH et par une faible expression d'une structure de ce genre ou même par son absence pour le locus Est-6. Ceci peut s'expliquer par la différence de types de sélection pour ces loci: la sélection dirigée pour l'allèle F de l'α-PGDH et la sélection stabilisante avec point d'équilibre ~0,45 pour l'allèle correspondant de l'Est-6. Il est clair que si la sélection stabilisante dépasse de beaucoup l'influence de la dérive génétique, elle masquera la structure de subpopulations et ses effets, déplaçant les concentrations d'allèles dans toutes les subpopulations vers la fréquence d'équilibre. Dans la sélection orientée vers la fixation de l'allèle F de l'α-phosphoglycérate-déshydrogénase, la contribution de la dérive génétique est apparemment plus sensible, ce qui conduit à l'établissement d'un équilibre de deux forces dont l'une est dirigée vers la diminution de la variabilité génétique du système et l'autre vers son augmentation. Cet équilibre réciproque se traduit par une apparition et un maintien prolongé d'une différenciation locale périodique des fréquences géniques, la migration des gènes étant d'une structure correspondante. Il est, de plus, important de souligner qu'une telle structure

et la fréquence génique moyenne qui lui correspond persistent même malgré la pression de sélection dirigée.

Pour conclure, les auteurs expriment leur reconnaissance à L. Jivotovski pour sa discussion des résultats et ses précieuses remarques.

RÉFÉRENCES

[1] *Алтухов Ю. П., Бернашевская А. Г., Милишников А. Н., Новикова Т. А.* Экспериментальное моделирование генетических процессов в популяционной системе *Drosophila melanogaster*, соответствующей кольцевой ступенчатой модели. Сообщение I. Обоснование подхода и особенности локальной дифференциации частот аллелей α-глицерофосфатдегидрогеназы и эстеразы-6. *Генетика*, 1979, 15, 4, 646.

[2] *Алтухов Ю. П., Победоносцева Е. Ю.* Экспериментальное моделирование генетических процессов в подразделенных популяциях. *Докл. АН СССР*, 1978, 238, 2, 466.

[3] *Алтухов Ю. П., Победоносцева Е. Ю.* Особенности генетического процесса в экспериментальной популяционной системе *Drosophila melanogaster*. *Ж. общ. Биол.*, 1979, 40, 6, 816.

[4] *Kimura M.* Stepping-stone model of population structure. *Ann. Rept. Nat. Inst. Genetics*, 1953, 3, 62.

[5] *Kimura M., Weiss G. H.* The stepping-stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, 1964, 49, 561.

[6] *Rohlf E. J., Schnell G. D.* An investigation of the isolation by distance model. *Amer. Natur.*, 1971, 105, 295.

[7] *Рычков Ю. Г., Шереметьева В. А.* Генетика циркумполярных популяций Евразии в связи с проблемой адаптации человека. В кн.: Ресурсы биосферы. 3. Л., Наука, 1976, 10.

[8] *Selander R. K.* Behavior and genetic variation in natural populations. *Amer. Zool.*, 1970, 10, 53.

MÉTHODE D'IDENTIFICATION DES PHÉNOTYPES MORPHOLOGIQUEMENT « MOYENS » ET « EXTRÊMES » D'APRÈS L'ENSEMBLE DES CARACTÈRES QUANTITATIFS

L. Jivotovski et Y. Altoukhov

Il existe beaucoup de preuves attestant que les plus stables en présence d'influences extrémales de l'environnement sont les individus qui se rapprochent de la moyenne de la population par l'ensemble de leurs caractères quantitatifs [3, 7]. Cette loi peut être interprétée comme le résultat de la « canalisation » du développement [13] due à l'action de la sélection stabilisante aux stades ontogénétiques précoces [11, 12]. Cependant, la question des génotypes que recèlent les phénotypes morphologiquement « moyens » et « extrêmes » n'est pas encore entièrement résolue. Du point de vue de la génétique des populations, la stabilisation du phénotype repose sur un système polygénique intégré, historiquement constitué dont l'importance est universelle [5]. Un rôle majeur dans sa formation est joué par l'hétérozygotie et la coadaptation des gènes [4, 9, 10]. Il faut admettre que les complexes adaptatifs de gènes formés par la sélection stabilisante se trouvent liés au phénotype « moyen optimal ». Les individus s'écartant de l'optimum se caractérisent par une faible adaptabilité.

La notion d'*optimum* (*norme*) est d'une importance essentielle en biologie et particulièrement pour la résolution de certains problèmes de la génétique des populations. Ceci soulève des questions relatives au caractère et à la nature des complexes adaptatifs de gènes. Actuellement, la voie la plus prometteuse, pour l'étude du problème, est la réunion des gènes mendéliens et des caractères quantitatifs dans la même recherche [1, 2, 6]. Grâce aux progrès de la génétique biochimique, l'étude de la structure du génotype pour de nombreux loci est un problème en principe résolu et n'a plus qu'un caractère purement technique. Il en est autrement pour l'élaboration de méthodes d'identification des phénotypes morphologiquement « moyens » et « extrêmes » d'après l'ensemble des caractères quantitatifs. Jusqu'à présent, ce problème se résolvait sans tenir compte des exigences de la statistique multidimensionnelle. Le but de la présente recherche est de combler, dans la mesure du possible, cette brèche.

© Издательство Наука, 1980.

1° **Désignations principales.** Admettons que l'on étudie p caractères quantitatifs. Chaque individu du phénotype x est caractérisé par un jeu (vecteur ligne) de valeurs de ces caractères: $x = (\xi_1, \xi_2, \ldots, \xi_p)$. La caractéristique principale de l'ensemble des caractères est le vecteur des valeurs moyennes $\bar{x}$, la matrice de covariances S et la matrice de corrélations R. En notation vectorielle-matricielle $\bar{x}$ et S sont déterminés par rapport à l'échantillon d'effectif N comme suit:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i \text{ et } S = \frac{1}{N-1} \left(\sum_{i=1}^{N} x_i^{\tau} x_i - N \bar{x}^{\tau} \bar{x} \right)$$

où τ est un symbole de la transposition, c'est-à-dire que x^{τ} est un vecteur colonne. Les éléments de la matrice de corrélations sont les coefficients de corrélation linéaire entre les caractères correspondants.

2° **Introduction de la métrique.** Pour résoudre le problème de l'identification des individus proches de la « norme adaptative », il convient de formaliser la notion de proximité et de déterminer la distance entre individus. Il est nécessaire d'introduire la métrique dans l'espace des caractères. Si les caractères étaient indépendants, il suffirait de donner une distance en métrique euclidienne ou en métrique équivalente avec normalisation préalable par rapport à l'écart type. Les caractères étant correlés, la métrique de l'espace p-dimensionnel des caractères est engendrée par la matrice non diagonale S. Déterminons la distance $d(x, y)$ entre les individus x et y de la sorte:

$$d(x, y) = \left[\frac{1}{p} (x - y) S^{-1} (x - y)^{\tau} \right]^{1/2}.$$

Par sa forme, cette distance est analogue à la distance entre populations de Mahalonobis [8], mais en diffère par le fait qu'ici on estime la distance entre phénotypes dans la population et, en outre, qu'elle est normalisée par rapport au nombre des caractères. Ce dernier point est important pour la comparaison des distances obtenues pour différents groupes de caractères.

Remarquons que la distance entre les individus d'une population donnée peut être considérée, d'une façon générale, en métrique engendrée par n'importe quelle matrice définie positive et non seulement en métrique d'une matrice de covariances obtenue pour cette population. A sa place, on pourrait prendre, par exemple, une matrice théorique fixée *a priori* ou une matrice estimée à partir d'autres données.

Vu la distance introduite, nous proposons un algorithme d'identification du groupe de phénotypes « moyens » M^0 et du groupe de phénotypes « extrêmes »: M^-, M^+ et M^0.

3° **Identification du groupe M^0.** Comme nous l'avons dit plus haut, la « norme adaptative » se trouve à proximité des valeurs moyennes des caractères. C'est pourquoi on peut prendre comme bonne approximation de cette norme la moyenne de la population $\bar{x}$. En vertu de la métrique introduite, la proximité de d_0 relativement à la moyenne pour un individu à phénotype x est trouvée à l'aide de la formule

$$d_0 = \left[\frac{1}{p}(x-\bar{x})\, S^{-1}\,(x-\bar{x})\right]^{1/2} .$$

La valeur minimale possible de d_0 est zéro. Elle est propre à un individu qui coïncide avec une moyenne de la population pour chacun de ses caractères, c'est-à-dire quand $x = \bar{x}$. Pratiquement, il n'exis-

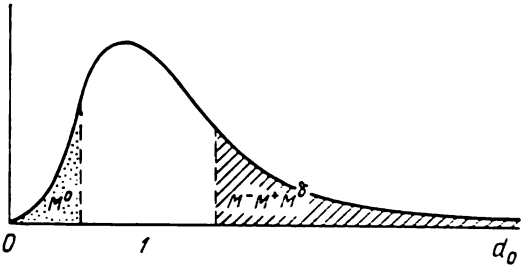


Fig. 1. Distribution des distances d_0 et identification des phénotypes morphologiquement «moyens» M^0 et « extrêmes » M^-, M^+, M^δ

te pas d'individus pour lesquels d_0 soit égale à zéro ou proche de zéro. Il y a également peu d'individus qui s'en écartent beaucoup, aussi la plupart des individus ont-ils une distance d_0 d'une valeur moyenne.

Si la distribution de l'ensemble des caractères est normale et si $\bar{x}$ et S sont des valeurs générales des paramètres, comme il découle de la théorie des distributions multidimensionnelles, la distance d_0 est distribuée comme $\frac{1}{\sqrt{p}}\chi^2_p$. Ici χ^2_p est une distribution χ^2 avec p degrés de liberté. La distribution de la distance d_0 pour $p > 2$ est représentée sur la figure 1. La moyenne de d_0^2 pour toute la population est égale à 1. C'est pourquoi les individus éloignés de $\bar{x}$ d'une distance de beaucoup inférieure à 1 peuvent être classés dans le groupe de phénotypes « moyens ». Nous figurerons ce groupe par le symbole M^0. La fraction des individus entrant dans M^0 sera désignée par m^0.

Pour une distribution normale des caractères, dans le groupe M^0 entrent les individus pour lesquels la valeur de d_0 ne dépasse pas

$\frac{1}{\sqrt{p}}\chi^2_p(m^0)$. Pratiquement, nous isolons le groupe M^0 de la façon suivante. Nous déterminons d'abord les valeurs échantillonnées de $\bar{x}$ et S et suivant ces dernières nous calculons pour chaque individu les distances d_0. Ensuite, nous rangeons tous les individus par ordre de distance croissante. En définitive, pour m^0 fixé, nous choisissons pour le groupe modal M^0 les individus qui sont les premiers par cet ordre en quantité $N^0 = m^0 N$.

Remarquons qu'on peut réaliser l'identification du groupe M^0 en ne se basant pas nécessairement sur les moyennes obtenues pour un échantillon donné. On peut prendre à leur place n'importe quel vecteur de caractères, théorique ou estimé à partir d'autres données.

4° **Identification des groupes M^-, M^+ et M^δ.** De même qu'on a identifié le groupe de phénotypes « moyens », on peut isoler le groupe d'individus les plus éloignés de la « norme ». Si la distribution des caractères est normale et si $\bar{x}$ et S sont des valeurs générales des paramètres, la fraction des individus ayant une distance supérieure à $\frac{1}{\sqrt{p}}\chi^2_p(1 - m^1)$ sera égale à m^1. Pratiquement, nous séparons ce groupe d'une rangée ordonnée suivant la distance en choisissant un nombre fixé ($N^1 = m^1 N$) d'individus ayant une valeur maximale de d_0.

Le groupe de phénotypes « extrêmes » n'a pas une composition homogène. Adressons-nous, par analogie, au cas d'un seul caractère: au groupe d'« extrêmes » appartiennent alors les individus ayant aussi bien de très petites que de très grandes valeurs du caractère. Dans ce cas, il y a une raison d'isoler le groupe de « petits » individus M^- et le groupe de « grands » individus M^+. Dans le cas de plusieurs caractères, nous trouvons parmi les individus « extrêmes », en plus de M^- et M^+, un groupe d'individus « disproportionnés » M^δ qui peuvent être rattachés à M^- pour certains caractères et à M^+ pour d'autres. Dans le cas d'un grand nombre de caractères, l'identification des trois groupes est liée à des difficultés du fait que les caractères différents peuvent varier dans des directions diverses. Cependant, pour le cas de caractères correlés, on peut proposer la méthode, pratiquement acceptable, de partition des groupes d'individus « extrêmes » en M^-, M^+, M^δ.

L'approche repose sur la méthode des composantes principales [8]. Pour l'identification des groupes, on observe l'ordre suivant des actions. Au début, d'après la matrice de corrélations on trouve le vecteur propre unitaire avec une valeur propre maximale λ_* — la première composante principale de la matrice R. Désignons-le par $x_* = (\xi_1^*, \xi_2^*, \ldots, \xi_p^*)$. Puis on procède à la normalisation par rapport aux écarts types correspondants. On cherche le vecteur $y_* = (\eta_1^*, \eta_2^*, \ldots, \eta_p^*)$ où $\eta_j^* = \xi_j^*/\sigma_j$ pour tous les $j = 1, \ldots, p$.

Après quoi on trouve pour chaque individu « extrême » la valeur de $d_1 = \frac{1}{\sqrt{p\lambda_*}} y_* (x - \bar{x})^\tau$ (pratiquement, pour faciliter les calculs, on peut se borner à trouver la valeur de $y_* x^\tau$ qui diffère de $d_1\sqrt{p\lambda_*}$ par la constante additive $y_* \bar{x}$).

Si la valeur de d_1 est de beaucoup inférieure à zéro (ou de beaucoup supérieure à zéro), ces individus seront rapportés au groupe M^- (ou au groupe M^+). Les autres individus « extrêmes » sont rattachés au groupe des « disproportionnés » M^δ. Ils ont les plus grandes valeurs du facteur de disproportionnalité que nous définissons comme $\delta = \sqrt{d_0^2 - d_1^2}$. Remarquons qu'en présence de corrélations négatives, l'interprétation de M^- (ou de M^+) en tant que groupe de « petits » (ou de « grands ») individus est conventionnelle.

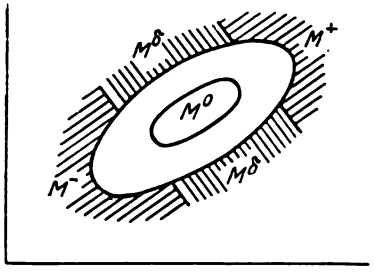


Fig. 2. Position des groupes M^0, M^-, M^+ et M^δ dans l'espace des caractères

Pratiquement, nous identifions les groupes indiqués de la façon suivante : nous commençons par ranger tous les individus « extrêmes » par valeurs de d_1, puis selon les coefficients de choix fixés m^-, m^+, m^δ ($m^- + m^+ + m^\delta = m^1$) nous rapportons les individus $N^- = m^- N$ au groupe M^-, les individus $N^\delta = m^\delta N$ au groupe M^δ, les individus $N^+ = m^+ N$ au groupe M^+. La partition de tout l'échantillon en groupes est présentée sur la figure 2 pour le cas de deux caractères.

CONCLUSION

Le procédé de calcul à deux échelons décrit ci-dessus permet d'identifier 4 groupes d'individus : le groupe d'individus « moyens » M^0, le groupe de « petits » individus M^-, le groupe de « grands » individus M^+ et le groupe d'individus « disproportionnés » M^δ. La génétique des populations a donc la possibilité de faire une analyse plus profonde des ensembles des caractères quantitatifs et polymorphes.

RÉFÉRENCES

[1] *Алтухов Ю. П., Ботвиньев О. К., Курбатова О. Л.* Популяционно-генетический подход к проблеме неспецифической биологической устойчивости человеческого организма. *Генетика*, 1979, 15, 2, 352-360.

[2] *Алтухов Ю. П., Животовский Л. А., Садыков С. С., Абдуллаев Б. А., Калабушкин Б. А.* Эффекты модального и направленного отбора по совокупности признаков у хлопчатника *Gossypium hirsutum L. Докл. АН СССР*, 1976, 227, 1, 212-215.
[3] *Bumpus H. C.* The elimination of the unfit as illustrated by the introduced sparrow. *Biol. Lect.*, 1899, 11, 209-226.
[4] *Dobzhansky Th.* Genetics of evolutionary process. N.Y.-L., Columbia Univ. Press, 1970.
[5] *Дубинин Н. П.* Экспериментальное исследование интеграции наследственных систем в процессах эволюции популяций. *Журн. общ. Биол.*, 1948, 9, 3, 203-244.
[6] *Дубинин Н. П., Алтухов Ю. П., Курбатова О. Л., Сусков И. И.* Интегральная генетическая характеристика « адаптивной нормы » в популяциях человека. *Докл. АН СССР*, 1976, 230, 4, 957-960.
[7] *Karn M. N., Penrose L. S.* Birth weight and gestation time in relation to maternal age, parity and infant survival. *Ann. Eugen.*, 1951, 15, 206.
[8] *Kendall M. G., Stuart A.* Design and analysis, and time-series. London, Charles Griffin, 1974.
[9] *Lerner M. J.* Genetic homeostasis. N. Y., J. Wiley, 1954.
[10] *Lewontin R. S.* The genetic basis of evolutionary change. N. Y.-L., Columbia Univ. Press, 1974.
[11] *Mayr E.* Animal species and evolution. Oxford, 1965.
[12] *Шмальгаузен И. И.* Факторы эволюции. Теория стабилизирующего отбора. М., Наука, 1968.
[13] *Waddington C. H.* The Strategy of genes. London, 1957.

EFFETS DE LA SÉLECTION MODALE ET DIRIGÉE POUR L'ENSEMBLE DES CARACTÈRES SUR LE COTONNIER *GOSSYPIUM HIRSUTUM L.*

Y. Altoukhov, L. Jivotovski, S. Sadykov, B. Abdoullaev et B. Kalabouchkine

Pour être résolu, le problème consistant à optimaliser les processus de sélection, d'amélioration et de création de races d'animaux et de variétés de plantes le mieux adaptées aux conditions du machinisme agricole, exige que de telles populations soient constituées d'individus suffisamment semblables par leur taille, leur forme, leur vitesse de développement et qu'elles possèdent une productivité élevée, résistent aux maladies et fournissent une production de belle qualité.

Comment faire pour que la race ou la variété répondent à tant d'exigences, parfois si contradictoires? La sélection dirigée ne permet pas toujours de résoudre ce problème: le renforcement de certaines fonctions entraîne inéluctablement un affaiblissement d'autres et une baisse de l'adaptabilité.

A cette approche on peut en opposer une autre, dans le cadre de laquelle on considère comme étant mieux adapté aux diverses exigences du milieu le phénotype qui s'écarte au minimum de la moyenne pour différents caractères variables [2, 3, 5, 7, 8, 10].

Le but de ce travail est de rechercher dans quelle mesure le principe de la sélection modale peut être efficace dans la culture cotonnière où il avait déjà été mis en pratique [4, 7, 9], mais ensuite rejeté comme inopérant [1]. Toutefois, il importe de souligner que les buts et les prémisses théoriques de notre recherche et des travaux cités sont différents.

Les expériences ont été effectuées en 1972-1974 à la ferme d'Etat « Bayaout-1 » de la région du Syr-Daria sur la parcelle d'essais de l'Institut de Biologie expérimentale des plantes de l'Académie des Sciences d'Ouzbékistan. On a choisi comme objet les variétés de cotonnier Tachkent, bien implantées dans la production. L'histoire de la création des variétés [6] montre leur hétérogénéité héréditaire, condition indispensable à tout travail de sélection.

© Издательство Наука, 1979.

Le schéma de l'analyse comprenait les étapes suivantes.

1° **Choix du matériel initial.** Le matériel a été recueilli sur des parcelles renfermant environ 5000 plantes des variétés Tachkent-1, Tachkent-2 et Tachkent-3. Les parcelles ont été conventionnellement divisées en carrés comprenant 4 rangées de 40 à 50 plants chacune. En choisissant un plant du centre de chaque rangée, nous avons sélectionné 100 plants de la variété Tachkent-1 et 40 plants de chacune des variétés Tachkent-2 et Tachkent-3. Ce matériel hétérogène représentait la génération parentale (*P*).

2° **Typologie morphologique des plantes et identification du phénotype « moyen ».** Chaque plante était caractérisée selon 17

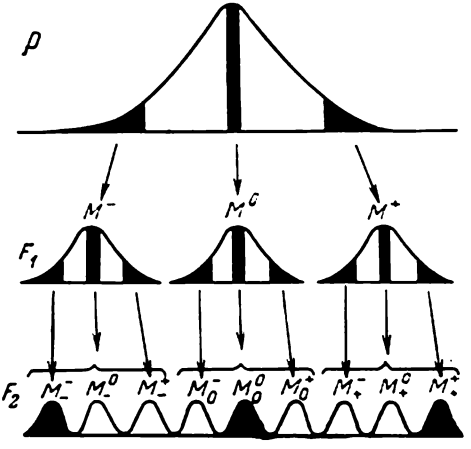


Fig. 1. Schéma de la sélection dans l'expérience:
Les groupes M^0 et M^0 sont formés par sélection modale. Les secteurs noircis correspondent aux types de groupes parentaux

traits liés à la morphologie et à la physiologie tels que hauteur, largeur, nombre de branches fruitières, nombre total de capsules, nombre de capsules ouvertes, taux des organes fruitiers tombés, etc.

On a considéré pour chaque caractère trois domaines de variabilité: un — proche de la moyenne, et deux — éloignés du centre de distribution de plus de deux écarts types vers la droite et vers la gauche. Les plantes qui, par la majorité de leurs caractères, appartenaient à l'un des domaines indiqués étaient respectivement rapportées aux groupes conventionnellement symbolisés par M^0 (phénotype « moyen »), M^+ et M^-.

3° **Analyse comparée des groupes de plantes soumises à la sélection modale et à la sélection dirigée.** Dans le but de tester les plantes par appréciation de leur descendance, on prélevait de la première et de la deuxième place sur les troisième et quatrième sympodes de

Particularités biologiques de trois populations de cotonnier obtenues culturales et de la

Groupes de plantes comparés	Nombre de plantes étudiées	Pouvoir germinatif des graines, %	Rendement par plante, g	Précocité, % des capsules ouvertes vers le 15 octobre
Témoin	152	93,3±2,1	106,6±10,6 (60)	88,4±2,6
M_{-}^{-}	114	96,6±1,6	72,7±7,2 (41)	97,8±1,3
M_{0}^{0}	644	98,0±0,5	101,2±5,4 (121)	93,6±0,9
M_{+}^{+}	130	92,5±2,2	130,1±12,7 (56)	87,8±2,8

Remarque. Dans les colonnes « rendement par plante », « longueur de la fibre » et

chaque buisson deux capsules. Les graines qu'elles contenaient étaient semées sur une parcelle d'essais dans un milieu à peu près identique. De la descendance de chacune des trois populations (F_1), on isolait de nouveau, en se basant sur les mêmes principes, les groupes respectifs de façon que F_2 fût déjà représenté par 9 sous-groupes (fig. 1). Trois d'entre eux — M_0^0, M_+^+, M_-^- — étaient comparés, en ce qui concerne la variabilité de tous les caractères étudiés, à la population témoin provenant des graines d'un semis de production.

Déjà, l'appréciation *de visu* des caractéristiques graphiques des groupes de plantes comparés montre l'effet de la sélection : dans le groupe M_0^0, relativement à toute une série de caractères, les écarts des plantes par rapport à la moyenne du témoin sont beaucoup moindres que dans les groupes M_+^+ et M_-^- (fig. 2). En effet, l'analyse statistique montre la haute certitude de ces différences ($P < 0{,}01$). La sélection modale a entraîné une diminution de la variabilité particulièrement nette dans la seconde génération (fig. 3).

Comparons maintenant les populations obtenues par sélection modale et dirigée pour d'autres caractères définissant leur productivité, la facilité de les récolter par des machines et leur résistance au wilt (*Tableau 1*). Nous voyons que pour toute une série de paramètres biologiques, le groupe de plantes à valeurs moyennes des caractères surpasse ou tout au moins n'est pas inférieur à la population témoin. La même tendance est évidente quand on compare le

Tableau 1

par sélection modale et dirigée tenant compte des exigences des façons cueillette mécanisées

Longueur de la fibre, mm	Poids de la fibre, g	Possibilité de mécaniser les travaux agricoles et la récolte, % des plantes de type optimal			Contamination par le wilt, %
		par la hauteur	par le niveau des capsules inférieures	par le caractère de la ramification	
32,7±0,23 (60)	2,02±0,20 (60)	44,7±4,0	50,6±4,0	62,5±3,9	10,9±2,4
33,0±0,80 (41)	2,09±0,11 (41)	35,1±4,5	50,0±4,7	68,4±4,4	8,8±2,7
33,6±0,20 (121)	2,34±0,06 (121)	59,3±1,9	61,8±1,9	78,6±1,7	3,6±0,8
32,2±0,36 (56)	2,33±0,15 (56)	28,5±3,9	41,5±4,3	50,0±4,3	18,4±3,4

« poids de la fibre », on indique entre parenthèses le nombre de capsules étudiées.

groupe M_0^0 et les groupes M_-^- et M_+^+. Les plantes M_+^+ se caractérisent par un rendement plus élevé. Cependant, il est clair que cet indice ne peut être rigoureusement estimé qu'au cours de l'épreuve de production. Compte tenu des données du *Tableau 1*, on voit que le rendement effectif de ces plantes sera fortement réduit si elles ne sont pas recoltables à la machine ou si elles résistent mal au wilt.

L'ensemble des données examinées permet de faire les conclusions suivantes *. 1° Dès la deuxième génération, la sélection modale d'après l'ensemble des caractères fait diminuer la variabilité dans la population et y augmenter la part des plantes proches du type « moyen » considéré comme optimal ; 2° Par certains caractères utiles à sa culture, surtout par la facilité de la mécanisation de sa récolte, et par sa résistance au wilt, cette population surpasse aussi bien le groupe témoin de plantes provenant de graines du semis de production que les deux autres populations formées au cours de la sélection dirigée.

Dans la discussion du génotype des plantes à phénotype optimal, il faut tenir compte de ce que dans une population se multipliant par croisement, elles peuvent receler un hétérozygote « multiple ». Il est clair qu'un tel phénotype ne peut rester stable dans une suite

* Etant donné qu'à cette étape de l'étude, les corrélations entre les caractères ne sont pas examinées, les conclusions sont purement qualitatives.

de générations assez longue et que sa reproduction n'est possible que par croisement de lignées homozygotes. Un tableau essentiellement différent doit être observé dans des populations autofécondées. Dans ce cas, le problème de la conservation durable du type optimal

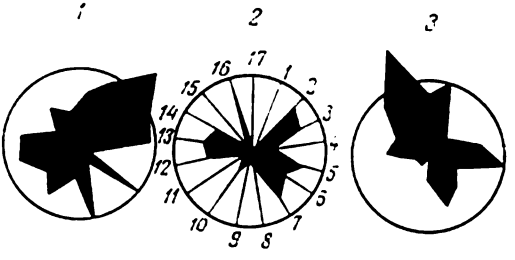


Fig. 2. Représentation graphique des écarts entre les moyennes dans les groupes M^{-}_{-} (*1*), M^{0}_{0} (*2*) et M^{+}_{+} (*3*) et les moyennes dans le témoin pour 17 caractères:
Sur les rayons on a porté les écarts par rapport à la moyenne du témoin

ou de types proches de celui-ci, s'ils sont homozygotes pour plusieurs gènes, peut être résolu sans efforts complémentaires.

On sait que le cotonnier est autogame. Cependant, la question de l'apport possible d'une pollinisation croisée à la formation de

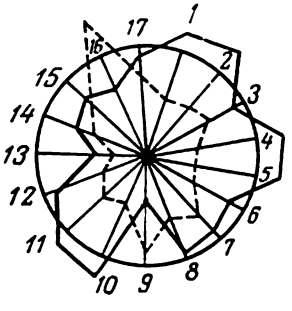


Fig. 3. Représentation graphique de la variabilité à l'intérieur des groupes M^{0} et M^{0}_{0} d'après 17 caractères:
Sur les rayons on a porté le rapport des variances des caractères dans ces groupes aux variances des caractères dans le témoin. La circonférence de rayon unitaire est le témoin, la ligne continue figure la première génération de la sélection modale (M^{0}) et la ligne pointillée, la deuxième génération de la sélection modale (M^{0}_{0})

diverses populations dans différentes conditions de l'environnement reste encore ouverte. Par conséquent, dans notre matériel les plantes ayant des moyennes des caractères choisis peuvent être aussi bien homozygotes qu'hétérozygotes et la tâche de l'avenir est d'étudier spécialement cette question.

Mais quoi qu'il en soit, les données déjà obtenues montrent l'efficacité de l'approche exposée par nous du problème consistant à optimiser le processus de sélection et à améliorer de cette façon les variétés de cotonnier d'une valeur économique. Cette conclusion est en contraste avec celle faite dans l'ouvrage [1] [v. également 4, 7, 9] affirmant l'inefficacité de la sélection modale pour le cotonnier.

Nous estimons que cette divergence est due aussi bien à des procédés méthodiques différents qu'au niveau d'hétérogénéité héréditaire du matériel initial.

RÉFÉRENCES

[1] *Arnold M. H.* Modal selection in BR-52. *Cotton Grow. Rev.*, 1972, 49, 2, 107-125.

[2] *Bumpus H. C.* The elimination of the unfit as illustrated by the introduced sparrow. *Biol. Lect.*, 1899, 11, 209-226.

[3] *Fisher R. A.* The Genetical theory of natural selection. Oxford, 1930.

[4] *Manning H. L.* Responce to selection for yield in cotton. *Cold. Spring Harb. Symp. Quant. Biol.*, 1955, 20, 103-110.

[5] *Mather K.* Polygenic inheritance and natural selection. *Biol. Rew.*, 1943, 18, 32-64.

[6] *Мирахмедов С. М., Юлдашев С. Х.* Вилтоустойчивые сорта хлопчатника. *Ташкент*, 1971, Ташкент.

[7] *O'Donald P.* Natural selection for quantitative characters. *Heredity*, 1971, 27, 1, 137-154.

[8] *Шмальгаузен И. И.* Факторы эволюции. Теория стабилизирующего отбора. М., 1968.

[9] *Walker J. T.* Modal selection in Upland cotton. *Heredity*, 1964, 19, 4, 559-583.

[10] *Wright S.* The genetical structure of populations. *Ann. Eugen.*, 1951, 15, 323-354.

PARTICULARITÉS DU DESSIN DU CALCUL ET STRUCTURE GÉNÉTIQUE DES GROUPES DE MOUTONS DE KARA-KOUL' RAPPORTÉS AUX TYPES MORPHOLOGIQUEMENT « MOYENS » ET « EXTRÊMES »

Y. Altoukhov, N. Sarsenbaev, K. Afanassiev, T. Malinina, A. Milichnikov et E. Pobédonostséva

Une des directions importantes du travail de sélection dans l'élevage des moutons de Kara-Koul' est le choix des progéniteurs afin d'obtenir une descendance ayant un type de caracul désiré. Bien que, comme on sait, les goûts changent avec le temps, le caracul type « jacket » est très estimé depuis de longues années par les connaisseurs [8, 9, 14, 19, 21, 22]. La valeur essentielle de ce caracul sont son dessin symétrique net et l'égalité de toutes les qualités requises pour estimer la fourrure (fig. 1).

La sélection visant à obtenir un type « jacket » a lieu depuis les années trente du XXe siècle [16]. Mais, au cours de la sélection directement d'après les caractères du caracul, un plateau a été atteint et les efforts ultérieurs sont restés inopérants [6]. Par exemple, même en 1975, la production d'agneaux à toison du type « jacket » au Kazakhstan n'a pas dépassé 36,6 % et la fréquence des animaux dont la fourrure était la plus estimée (« jacket-1 ») n'était que de 2 %. Un tableau de ce genre est apparemment observé du fait que le développement des principales particularités morphologiques du type de caracul dépend des effets produits par l'interaction de nombreux gènes dans le cadre d'un système génétique suffisamment hétérozygote, rendant l'ontogénèse dans son ensemble « canalisée » et « tamponnée » [10, 15, 23]. De ce point de vue, on peut postuler la connexion suivante: plus l'ontogénèse prénatale résiste aux différentes fluctuations du milieu tant extérieur qu'intérieur, plus il y a de chances d'obtenir une descendance présentant le moins possible d'écarts des différents systèmes morphofonctionnels de l'organisme, y compris les particularités structurales de la toison caractérisées avant tout par une symétrie déterminée.

Nous avons montré dans nos recherches précédentes qu'une telle stabilité non spécifique est propre au phénotype « moyen » (modal) surtout, si la typologie repose sur un nombre suffisamment grand de caractères anatomo morphologiques polygéniques et si la variance

© Издательство Наука, 1980.

génétique (hétérogénéité) de la population est suffisamment élevée [2, 5]. Des résultats très proches ont été obtenus également grâce aux travaux d'autres chercheurs [7, 20].

A la lumière de ces données, il devient évident que si nous voulons augmenter parmi la descendance la part des animaux ayant

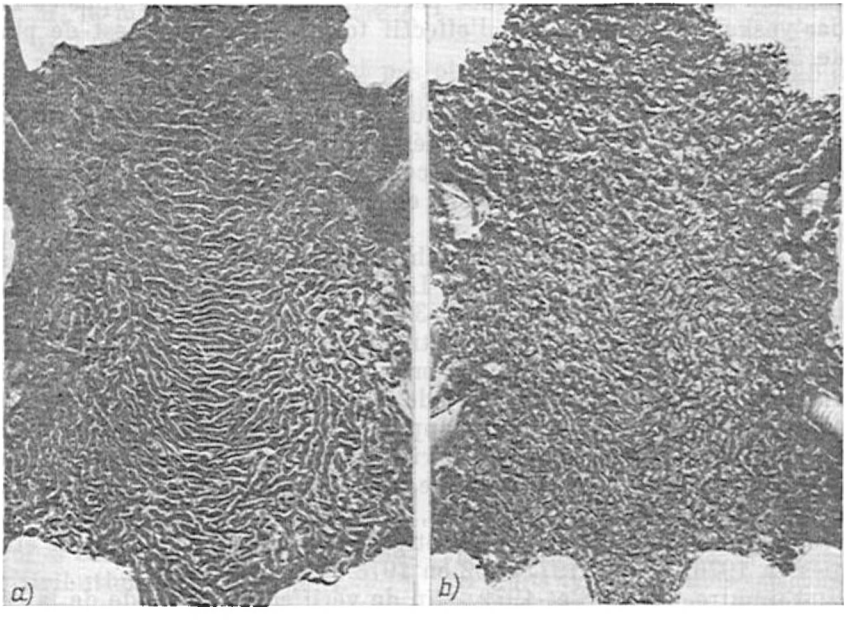


Fig. 1. Types désiré (*a*) et non désiré (*b*) de caracul:
a — type « jacket »: boucles moyennes, longues côtes semi-circulaires, dessin net, concentrique et symétrique, beaux lustre et soyeux, peau suffisamment fine; *b* — type caucasien: boucles trop grosses, côtes courtes, « haricot » grossier, soyeux et lustre très faibles, dessin irrégulier

un dessin de caracul qui nous intéresse, nous devons choisir les parents de façon que la descendance soit, *grosso modo*, moyenne pour ce qui concerne les particularités constitutionnelles.

Le but de notre recherche est de tester cette hypothèse.

MATÉRIEL ET MÉTHODE

Le matériel a été recueilli en 1976-1979 à la Ferme de sélection d'Etat « Zadar'ynskaïa » du district Bougoun de la région de Tchimkent, entreprise expérimentale de l'Institut de recherche sur l'élevage des moutons de Kara-Koul' du Kazakhstan. Depuis son organisation (1940), l'entreprise se spécialise dans l'élevage d'agneaux

caraculs noirs donnant une toison type « jacket ». Elle est composée d'un troupeau provenant de cinq fermes d'Etat d'Ouzbékistan et de quatre fermes d'Etat de Turkménie. Pour les buts que se propose notre recherche, c'est un fait significatif, car il révèle l'hétérogénéité héréditaire marquée du troupeau. Au cours de toute l'existence de l'entreprise, on s'y livre à un travail de sélection ayant pour résultat la création d'une race particulière de moutons appelée zadar'ynskaïa. Actuellement, l'effectif total du troupeau est de plus de 50 000 animaux.

Bien que, dans l'entreprise, le rendement du caracul des meilleures qualités soit assez élevé, nous nous heurtons également ici au même problème : l'absence de schéma suffisamment justifié de choix des parents pour obtenir une descendance ayant un type de caracul optimal. Au total, le rendement en caracul « jacket-1 » ne dépasse pas 16 %.

Le schéma de notre recherche incluait les étapes suivantes.

1° **Choix du matériel initial.** L'échantillon a été tiré d'un grand troupeau d'un effectif d'environ 3000 moutons. Afin d'embrasser le plus entièrement possible le patrimoine génétique collectif de la population, les agneaux étaient pris au hasard après avoir été individuellement appréciés par ordre de qualité. Immédiatement après sa naissance, chaque agneau était apprécié par l'ensemble de caractères reflétant aussi bien ses particularités constitutionnelles (poids, longueur du corps, tour de poitrine, etc.) que la qualité du caracul (longueur et largeur de la boucle, longueur du poil, épaisseur de la peau, type du dessin, etc.). De cette façon, 350 agneaux ont été étudiés en 1976, 618 en 1977, 202 en 1978 et 161 en 1979.

En outre, en 1978 et 1979, afin de vérifier l'exactitude de la détermination des races et d'étudier la structure de la population d'après les gènes mendéliens, on identifiait par électrophorèse en gel de polyacrylamide les génotypes de 19 béliers, de 101 brebis et de 256 agneaux de race pour les loci de la transferrine (sidérophiline), de l'hémoglobine et de l'estérase sérique. 215 prises de sang ont été analysées directement dans la ferme de sélection et 161 à notre laboratoire (le matériel ayant été transporté du Kazakhstan à Moscou dans des bouteilles isolantes avec de la glace).

L'électrophorèse était opérée suivant les méthodes standards admises dans notre laboratoire [4] avec certaines modifications.

2° **Typologie morphologique des animaux et identification des types « moyens » et « extrêmes ».** L'identification des individus « moyens » et « extrêmes » dans l'échantillon était faite simultanément selon trois caractères corrélés déterminant les particularités constitutionnelles de l'animal (*Tableau 1*). Ont été utilisées les approches exposées plus haut [3, 5, 11] et permettant, par la somme des écarts types, d'estimer le degré de distinction de chaque animal

par rapport à la moyenne de la population:

$$\sum_{i=1}^{3} \frac{V_i - M_i}{\sigma_i}$$

où V_i est la valeur individuelle du caractère i dans l'échantillon,
M_i la moyenne du caractère i,
σ_i l'écart type.

Les individus ne s'écartant pas plus de 1,0 σ à droite et de 0,5 σ à gauche du centre de distribution pour chaque caractère étaient considérés comme « moyens » (fig. 2). Nous avons symbolisé ce groupe par M^0 (« norme adaptative »). Il est représenté par des agneaux ayant les valeurs suivantes des caractères (v. *Tableau 1*): 1° de 4100 à 4800 g; 2° de 31 à 34 cm; 3° de 37 à 40 cm. Les valeurs des caractères pour les groupes M^+ et M^- étaient respectivement 1° > 5100 et < 3600 g; 2° > 36 et < 29 cm; 3° > 41 et < 35 cm. Ceci signifie que les groupes M^+ et M^- embrassaient les animaux qui s'écartaient du centre de distribution de 3 σ et plus.

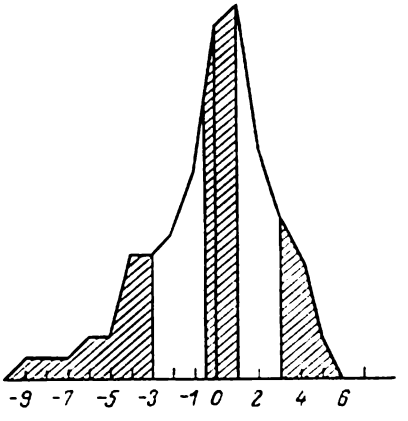


Fig. 2. Distribution de la somme des écarts types, par trois caractères polygéniques, dans un échantillon d'agneaux caraculs nouveau-nés ($N = 618$):

Sur l'horizontale, $\sum_{i=1}^{3} \frac{V_i - M_i}{\delta_i}$; sur la verticale, la fréquence des animaux. Les autres explications sont présentées dans le texte

Etant donné que cette méthode de classification peut conduire à des erreurs rattachées tant à l'effet de « centrage », surtout dans le cas de caractères non corrélés ou peu corrélés, nous avons fait une analyse complémentaire de tout le matériel et exclu de l'examen, à une étape donnée du travail, les animaux présentant des combinaisons « discordantes » de caractères. En d'autres termes, nous analysons ici seulement les animaux qui ont des types biologiquement « moyens » ou « extrêmes » pour l'ensemble de ces trois caractères.

Après une telle typologie, le groupe moyen (« norme adaptative ») M^0 comportait 216 animaux et les groupes M^- et M^+ renfermaient respectivement 101 et 97 individus (v. fig. 2). Par la suite, le mode d'identification des phénotypes « moyens » et « extrêmes » a été perfectionné compte tenu des exigences de la statistique multidi-

Tableau 1

Variabilité des agneaux caraculs nouveau-nés suivant trois caractères quantitatifs ($N = 618$)

Caractère étudié	Paramètres des distributions			Coefficient de corrélation
	M	m	σ	
1° Poids à la naissance, g	4,395	0,03	0,69	$r_{1,2} = 0,83 \pm 0,007$
2° Longueur oblique du corps, cm	32,31	0,08	2,00	$r_{1,3} = 0,90 \pm 0,004$
3° Tour de poitrine, cm	38,11	0,10	2,43	$r_{2,3} = 0,83 \pm 0,007$

mensionnelle et a été effectué sur un miniordinateur « Hewlett-Packard 9830 » (U.S.A.). On n'a pas découvert de divergences significatives des résultats de la classification.

3° **Estimation de la qualité du caracul.** En accord avec les procédés élaborés par plusieurs chercheurs étudiant les questions relatives à la qualité du caracul [8, 9, 14, 19, 22] et avec l'instruction d'appréciation par ordre de qualité des agneaux caraculs [13], nous avons établi les indices suivants caractérisant le type de caracul optimal:

Type de caracul: « jacket »
Longueur de la boucle, mm: 30,0 et plus
Largeur de la boucle, mm: 4,0 à 8,0
Longueur du poil, mm: pas plus de 10
Epaisseur de la peau, mm: 1,8 à 2,5
Soyeux de la fourrure: prononcé, normal
Lustre de la fourrure: beau, normal.

En se basant sur ces indices, nous avons pu ensuite déterminer la fréquence des animaux dont la fourrure caracul avait les qualités les plus précieuses parmi les trois groupes d'agneaux (M^-, M^0, M^+) isolés uniquement suivant leurs particularités constitutionnelles.

4° **Electrophorèse des protéines sanguines.** Les prises de sang de 2 ml étaient prélevées sur la jugulaire et centrifugées pendant 20 mn à 7000 t/mn. Les hématies étaient hémolysées en alternant la réfrigération et le dégel. On ajoutait de l'eau distillée refroidie en proportion de 1 : 1 et l'on centrifugeait les hémolysats pendant 20 mn à 15 000 t/mn.

L'électrophorèse était effectuée dans un bloc vertical de gel de polyacrylamide à concentration de 5,5 et 7,5 %. Système tampon THAM*-EDTA**-borate [1, 11, 17]. Pour colorer les zones de trans-

ferrine (sidérophiline), on utilisait une solution à 0,025% de Koumassi bleu en acide triacétique. Les phénotypes de l'hémoglobine sont faciles à identifier sur les électrophorégrammes frais sans addition de colorant. Pour mettre en évidence l'activité estérasique, on utilisait la réaction histochimique à l'α-naphtylacétate.

PARTIE EXPÉRIMENTALE

Comme nous l'avons indiqué plus haut, une des tâches essentielles de ce travail est de déterminer dans quelle mesure le type de caracul qui intéresse l'exploitant se rattache aux particularités constitutionnelles de l'animal. Etant donné que des caractères quantitatifs, tels que la taille et le poids, impliquent une composante paratypique, nous avons construit notre classification en tenant compte de la variabilité de ces caractères chez les agneaux nouveaunés, c'est-à-dire lorsque les différences individuelles déterminées par le génotype se manifestent plus nettement qu'aux stades ontogénétiques ultérieurs. Si l'hypothèse d'une connexion non spécifique entre la stabilité générale de l'ontogénèse et le développement du type de caracul optimal est juste, cela doit se manifester aussi avec une grande netteté aux stades ontogénétiques précoces.

Tableau 2

Fréquence (%) des différents types de caracul dans les trois groupes d'agneaux classés suivant leurs particularités constitutionnelles

Type de caracul	Type de constitution			Certitude des différences (F_φ) de la fréquence du type « jacket »	
	M^- ($N = 101$)	M^0 ($N = 216$)	M^+ ($N = 97$)		
« Jacket »	18,80	43,05	22,70		
Côtelé	5,90	2,80	8,20	M^-, M^+	2,18
Plat	1,98	3,24	4,12	M^-, M^0	18,7 *
Caucasien	73,32	50,91	64,98	M^+, M^0	6,5 **
Total	100	100	100		

* $P < 0,01$.
** $P < 0,001$.

En effet, les données du *Tableau 2* nous persuadent de la justesse de cette conjecture : la distribution des fréquences de différents types de caracul n'est pas la même parmi les trois groupes d'individus classés suivant leurs particularités constitutionnelles : dans le groupe M^0 la fréquence des animaux type « jacket » dont la frisure est paral-

* THAM ou TRIS : *tris*-(hydroxyméthyl) aminométhane. (*N.d.R.*)
** Ethylènediamine tétraacétate. (*N.d.R.*)

lèle et concentrique, atteint 43,05 % contre respectivement 18,8 et 22,7 % dans les groupes M^- et M^+. Parmi les animaux du groupe M^0 la part des agneaux à fourrure de qualité « jacket-1 » est de 28,2 %. On voit que d'après les qualités du caracul les animaux « moyens »

Tableau 3

Indices de qualité du caracul dans trois groupes d'agneaux se distinguant par les particularités constitutionnelles

Groupe d'animaux	Effectif de l'échantillon	Particularités de la fourrure caracul							
		Longueur de la boucle, mm		Largeur de la boucle, mm		Longueur du poil, mm		Epaisseur de la peau, mm	
		M	*m*	*M*	*m*	*M*	*m*	*M*	*m*
M^-	101	39,33	1,69	4,88	0,09	9,30	0,12	1,69	0,03
M^0	216	47,70	1,43	5,19	0,05	9,54	0,07	2,10	0,02
M^+	97	48,13	2,64	5,48	0,07	10,15	0,13	2,69	0,05

sont également proches du type optimal tant en ce qui concerne la valeur des caractères (*Tableau 3*) que l'ampleur de leur variation: l'égalisation maximale des indices tels que la longueur du poil, la longueur et la largeur de la boucle est typique des animaux du groupe M^0 (fig. 3).

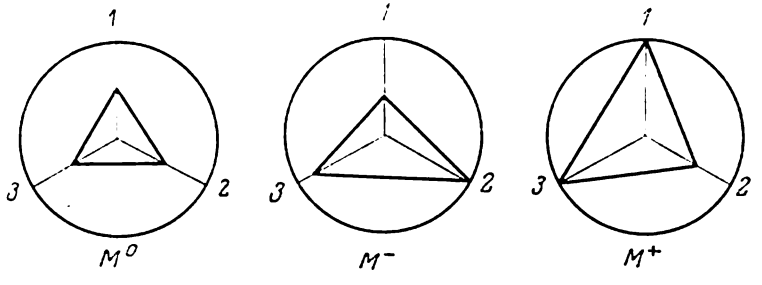


Fig. 3. Confrontation graphique des coefficients de variation de trois caractères quantitatifs, définissant la qualité du caracul dans les groupes d'agneaux M^0, M^-, M^+:

La valeur maximale de la variabilité de chaque caractère dans le groupe est posée égale à 100 % et correspond au rayon du cercle: *1* — longueur de la boucle; *2* — largeur de la boucle; *3* — longueur du poil

Examinons maintenant les particularités de la structure génétique du groupe morphologiquement « moyen » et de deux groupes « extrêmes » d'agneaux pour les loci de la transferrine (sidérophiline), de l'hémoglobine et de l'estérase du sang.

Le polymorphisme héréditaire de la transferrine (sidérophiline) sérique et de l'hémoglobine des moutons caraculs est bien connu d'après les travaux de nombreux auteurs [12] et son interprétation génétique ne suscite pas de difficultés : pour un des loci hémoglobiniques on connaît deux principaux allèles codominants β^A et β^B et trois génotypes — deux homozygotes Hb AA et Hb BB et un hétérozygote Hb AB. Le locus Tf est polyallélique à un type d'hérédité

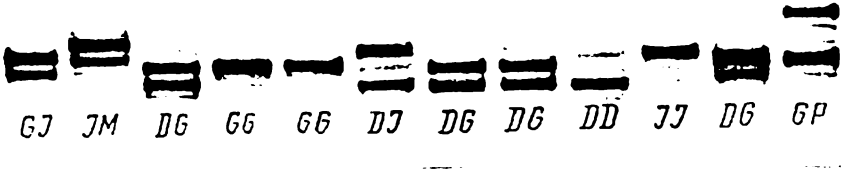


Fig. 4. Electrophorégramme des types de transferrine découverts chez les moutons de Kara-Koul' :
Electrophorèse en GPA, solution tampon THAM-EDTA-borate, pH 8,6

codominant. Dans les échantillons que nous avons étudiés, on a découvert 6 allèles, 16 génotypes pour la transferrine (fig. 4) et 2 génotypes (2 allèles) pour l'hémoglobine.

Des données sur le polymorphisme génétique des arylestérases sérique et érythrocytaire sont présentées dans un travail récapitulatif de Rich et coauteurs [18]. Dans nos expériences à l'aide de l'électrophorèse sur gel de polyacrylamide *, nous avons mis en évidence, dans le sérum sanguin, trois variantes d'estérase chez les animaux adultes et jusqu'à cinq variantes chez les agneaux (fig. 5). Il est supposé que, de même que pour les autres systèmes de polymorphisme biochimique, nous avons affaire ici à une simple hérédité codominante au moins de quatre allèles d'un locus. Dorénavant, jusqu'à un déchiffrement génétique plus détaillé de ce polymorphisme, nous nous en tiendrons au schéma proposé.

Les particularités génétiques des groupes d'animaux comparés sont présentées aux *Tableaux 4-6* établis de façon à pouvoir juger du nombre de génotypes observé et calculé en vertu de la formule de Hardy-Weinberg (entre parenthèses) ainsi que des fréquences géniques et des valeurs de χ^2. Les niveaux d'hétérozygotie sont figurés par la grandeur D qui représente le rapport de la différence entre le nombre d'hétérozygotes observé et celui calculé à leur nombre calculé.

* Bien que l'EDTA (éthylènediamine tétraacétate) inhibe l'arylestérase, nous n'avons pas observé dans nos expériences d'affaiblissement de la couleur des zymogrammes quoique on ait utilisé le système tampon THAM-EDTA-borate.

Tableau 4

Distribution des génotypes et des fréquences d'allèles de l'hémoglobine dans trois groupes comparés d'agneaux caraculs

Groupe d'animaux	Effectif de l'échantillon	Génotype			Fréquence du gène ($P \pm SE$)		χ^2	D
		AA	AB	BB	A	B		
M^-	64	(0,01)	1 (0,99)	63 (63,00)	0,0078± ±0,0078	0,9922± ±0,0078	0,007	0,006
M^0	67	(0,06)	4 (3,89)	63 (63,05)	0,0299± ±0,0147	0,9701± ±0,0147	0,0636	0,0293
M^+	74	(0,05)	4 (3,89)	70 (70,06)	0,0270± ±0,0133	0,9730± ±0,0133	0,0673	0,0285

Remarque. Entre parenthèses on a donné, de même que dans les tableaux suivants, les fréquences des génotypes calculées d'après la formule de Hardy-Weinberg.

L'examen de ces données montre les tendances suivantes.

1° Le niveau d'hétérozygotie pour le locus Hb est plus élevé dans les groupes M^0 et M^+ génétiquement identiques pour ce caractère que dans le groupe M^-.

2° Le même tableau est observé aussi pour le locus transferrine bien que le groupe M^+ soit plus hétérozygote que le groupe M^0.

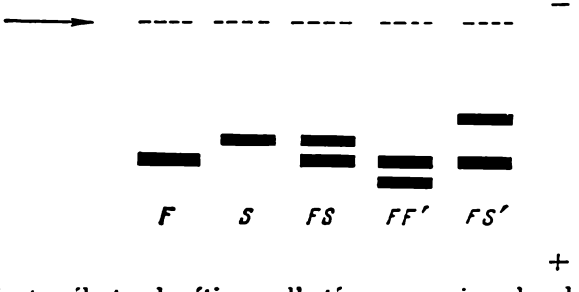


Fig. 5. Variantes électrophorétiques d'estérase sanguine chez les moutons de Kara-Koul':
La flèche indique la position de départ

3° Pour le locus estérase les groupes M^- et M^+ se trouvent un peu plus hétérozygotes que le groupe M^0.

4° Des génotypes rares sont rencontrés plus souvent parmi les animaux se rapportant aux types morphologiquement « extrêmes » M^+ et M^-. Ceci se remarque particulièrement pour les génotypes de la transferrine (DD, JS et MS) et de l'estérase (FS′, SS). Un allèle rare TfS se rencontre seulement dans les groupes M^+ et M^-.

Tableau 5

Distribution des génotypes et des fréquences des allèles de la transferrine dans trois groupes comparés d'agneaux caraculs

Groupe d'agneaux	Effectif de l'échantillon	Génotype								
		DD	DG	DJ	DM	DP	GG	GJ	GM	CP
M⁻	64	1 (2,69)	9 (10,00)	8 (12,31)	2 (4,36)	(0,51)	6 (9,28)	15 (22,85)	2 (8,09)	1 (0,95)
M°	67	5 (0,30)	3 (2,75)	1 (3,70)	1 (1,55)	(0,40)	5 (6,27)	19 (16,83)	3 (7,04)	4 (1,84)
M⁺	74	2 (0,57)	9 (4,21)	2 (5,54)	(1,49)	(0,53)	9 (7,78)	27 (20,43)	(5,51)	1 (1,95)

Groupe d'agneaux	Effectif de l'échantillon	Génotype							
		JJ	JM	JP	JS	MM	MP	MS	DS, PS, GS, SS, PP
M⁻	64	6 (14,06)	11 (9,96)	1 (1,17)	1 (0,59)	1 (1,76)	(0,42)	(0,21)	(0,79)
M°	67	8 (11,28)	16 (9,44)	1 (2,47)		1 (1,97)	1 (1,03)		(0,13)
M⁺	74	4 (13,41)	16 (7,24)	3 (2,55)	(0,43)	(0,98)	(0,69)	1 (0,12)	(0,57)

Groupe d'agneaux	Effectif de l'échantillon	Fréquence du gène ($P \pm SE$)						χ^2	D
		D	G	J	M	P	S		
M⁻	64	0,1641± ±0,0327	0,3047± ±0,0407	0,3750± ±0,0428	0,1328± ±0,0300	0,0156± ±0,0110	0,0078± ±0,0078	11,8926	0,0824
M°	67	0,0672± ±0,0216	0,3060± ±0,0398	0,4104± ±0,0425	0,1716± ±0,0326	0,0448± ±0,0179		15,2616	0,1266
M⁺	74	0,0878± ±0,0233	0,3243± ±0,0385	0,4257± ±0,0407	0,1149± ±0,0262	0,0405± ±0,0162	0,0068± ±0,0068	44,3331	0,1927

Tableau 6

Distribution des génotypes et des fréquences des allèles de l'estérase dans trois groupes comparés d'agneaux caraculs

Groupe d'agneaux	Effectif de l'échantillon	Génotype								
		FF	FS	FF′	FS′	SS	SF′	SS′	F′F′	S′S′
M^-	64	31	27		3	3				
		(33,07)	(23,72)		(2,15)	(4,26)		(0,77)		(0,03)
M^0	67	37	27	1		2				
		(38,82)	(23,59)	(0,76)		(3,59)	(0,23)		0,01	
M^+	74	30	36	1	1	6				
		(32,45)	(31,79)	(0,66)	(0,66)	(7,78)	(0,33)	(0,33)	(0,00)	(0,00)

Groupe d'agneaux	Effectif de l'échantillon	Fréquence du gène ($P \pm SE$)				χ^2	D
		F	S	F′	S′		
M^-	64	0,7188±0,0397	0,2578±0,0387		0,0234±0,0134	2,0962	0,1261
M^0	67	0,7612±0,0368	0,2313±0,0364	0,0075±0,0075		1,5926	0,0981
M^+	74	0,6622±0,0389	0,3244±0,0385	0,0067±0,0067	0,0067±0,0067	2,1640	0,1254

5° La comparaison des groupes d'agneaux morphologiquement « moyens » et « extrêmes » d'après les trois loci protéiques montre que les niveaux minimal et maximal de l'hétérozygotie totale sont respectivement propres aux groupes M^- et M^+, alors que dans le groupe M^0 l'hétérozygotie s'avère moyenne. Cependant, il n'y a de raison de parler de différences de ce genre que comme d'une tendance, étant donné que pour obtenir une estimation stable du niveau d'hétérozygotie le jeu de loci doit être augmenté.

Pourtant, il est évident que de rares combinaisons de génotypes doivent avoir une adaptabilité moindre aux stades ontogénétiques précoces. De plus, la pression de sélection négative est la plus forte sur le groupe « extrême » M^- dans lequel la fréquence des génotypes rares est notablement plus élevée que parmi les animaux morphologiquement « moyens » (*Tableau 7*).

Tableau 7

Fréquences des génotypes dans trois groupes d'agneaux

Groupe d'agneaux	Effectif de l'échantillon	Combinaisons de génotypes pour les loci Hb, Tf, Est			F_φ
		fréquentes (> 0,12)	moyennes (0,04 à 0,12)	rares (< 0,04)	
M^0	67	0,1642	0,3721	0,4627	M^0, M^- 5,00 *
M^-	64	0,1406	0,2031	0,6553	M^0, M^+ 0,35
M^+	74	0,1486	0,3378	0,5135	M^-, M^+ 2,92

* $P < 0,01$.

En même temps les génotypes moyens pour la fréquence de leur occurrence prédominent dans le groupe moyen M^0.

Ainsi, les données que nous avons obtenues établissent une connexion évidente entre le type de caracul, les particularités constitutionnelles des moutons de Kara-Koul' et leur structure génétique pour les loci étudiés. Par conséquent, il devient possible de mieux choisir les parents pour obtenir une descendance ayant un type de caracul désiré. Il est éviqent que les croisements de types morphologiquement « extrêmes » ou proches d'eux peuvent présenter un intérêt particulier. De plus, il serait important d'analyser les différentes variantes de croisements et d'utiliser, pour l'identification précoce des individus les plus prometteurs, le jeu le plus large possible de caractères quantitatifs déterminant les particularités constitutionnelles de l'animal. Malheureusement, il n'est pas si simple d'effectuer ce travail : dans la ferme de sélection « Zadar'ynskaïa » les béliers reproducteurs

appartiennent principalement au type M^+ par les particularités de leur constitution. En outre, comme nous l'avons constaté, les erreurs de détermination des races atteignent au moins 40 %.

Cependant, les résultats des croisements présentés au *Tableau 8*

Tableau 8

Fréquence des agneaux caraculs de types morphologiquement « moyens » (M^0) et « extrêmes » (M^+ et M^-) dans la descendance des mères correspondantes

Mères	Descendance		
	M^0	M^-	M^+
M^-	0,61±0,10	0,04±0,04	0,35±0,10
M^0	0,38±0,07	0,10±0,04	0,52±0,08
M^+	0,43±0,09	0,20±0,07	0,37±0,09

se sont avérés très prometteurs. Comme on s'y attendait, la fréquence maximale des agneaux morphologiquement « moyens » a été découverte parmi les descendants des mères du groupe M^-.

Les effectifs des groupes comparés ne sont pas assez grands pour attester une certitude statistique élevée de ces données. Pourtant, les tendances relevées sont si évidentes que l'on peut avec beaucoup d'assurance s'attendre à les voir confirmées par l'analyse d'un matériel abondant.

DISCUSSION

La connexion, découverte dans ce travail, entre les particularités constitutionnelles des agneaux caraculs et le type de leur toison n'est pas inattendue, l'analyse des publications sur l'élevage des moutons de Kara-Koul' montre qu'une corrélation de ce genre avait déjà été, sous telle ou telle forme, signalée auparavant [14]. Cependant, jusqu'à présent, on n'avait pas essayé d'analyser cette question à partir des positions de la génétique des populations, surtout en tenant compte de la connexion évidente entre la structure intégrale du génotype et la stabilité de l'ontogénèse. Un examen de ce genre révèle que le type morphologiquement « moyen » est caractérisé non seulement par une grande viabilité, mais aussi par une vaste stabilité non spécifique des stades ontogénétiques précoces. Apparemment, le complexe coadapté de gènes correspondant assure le contrôle du développement optimal et harmonieux de tout système morphofonctionnel dans la caractéristique duquel les propriétés de symétrie et

de « normalisation » dans l'expression des caractères occupent une place importante.

Cependant, une autre chose est compréhensible. Toute la stratégie de la sélection est ordinairement orientée vers l'amélioration des différentes qualités ou propriétés de la race ou de la variété, ce qui automatiquement, par tradition, s'associe à la sélection des variantes « optimales ». Il est clair qu'une telle voie n'est pas toujours justifiée : au cours de la sélection dirigée l'hypertrophie de certains caractères et qualités est inévitable en raison de l'existence de corrélations négatives dans le système de l'ontogénèse, elle se conjugue parfois à un affaiblissement et à une détérioration d'autres caractères — « faux frais » de la sélection. Mais cette approche est tout à fait inacceptable lorsque la notion de meilleur et d'optimal coïncide avec la notion ou le type « moyen ». C'est précisément à une telle situation que nous nous heurtons dans le cas des particularités du dessin de la toison qui nous intéressent chez les agneaux caraculs ou bien dans le cas de races d'animaux ou de variétés de plantes se distinguant par leur stabilité maximale et par leur adaptation aux méthodes d'exploitation industrielles [2, 5].

Ainsi, malgré les résultats univoques obtenus dans ce travail, les auteurs n'estiment pas avoir trouvé la solution définitive du problème du choix de progéniteurs ne produisant que des animaux à type de caracul désiré. Afin de résoudre ce problème, il est nécessaire d'élargir notablement les recherches correspondantes et avant tout

1° d'augmenter le jeu de caractères quantitatifs utilisés pour l'identification des groupes d'individus « moyens » et « extrêmes »,

2° de mettre au point un système de classification des types par les caractères anatomomorphologiques compte tenu du sexe des animaux, estimant que pour les mâles et les femelles les limites de la « norme adaptative » diffèrent,

3° d'augmenter le jeu de marqueurs génétiques polymorphes (protéines et groupes sanguins) afin que le choix des parents suivant leurs particularités constitutionnelles s'associe à celui porté sur l'appréciation du génotype d'après l'ensemble des gènes mendéliens,

4° de vérifier sur la descendance le nombre le plus grand possible de combinaisons de progéniteurs se distinguant par les particularités constitutionnelles et la structure intégrale du génotype.

Si l'on parvient à réaliser ce programme avec succès, on peut espérer que l'approche justifiée dans ce travail ouvre une voie nouvelle à la cause d'une meilleure efficacité de la sélection en vue de la qualité du caracul. Il est évident que la réalisation de ce modèle exigera la révision de certains principes de sélection bien ancrés.

Les auteurs profitent de l'occasion pour remercier G. Popov, M. Chirinski, J. Glembotski, I. Diatchkov et M. Rich de leurs précieuses consultations.

RÉFÉRENCES

[1] *Алтухов Ю. П.* Популяционная генетика рыб. М., Пищевая промышленность, 1974.

[2] *Алтухов Ю. П., Абдуллаев Б. А., Садыков С. С.* О возможности использования принципа модальной селекции для стабилизации и улучшения сортов хлопчатника. В сб.: Генетика и селекция количественных признаков хлопчатника. Ташкент, ФАН, 1978, 19.

[3] *Алтухов Ю. П., Ботвиньев О. К., Курбатова О. Л.* Популяционно-генетический подход к проблеме неспецифической биологической устойчивости человеческого организма. Сообщение I. Постановка проблемы и обоснование подхода. Параметры распределений антропометрических признаков новорожденных и грудных детей в норме и патологии. *Генетика*, 1979, 15, 2, 352.

[4] *Алтухов Ю. П., Салменкова Е. А., Омельченко В. Т., Сачко Г. Д., Слынько В. И.* О числе мономорфных и полиморфных локусов в популяции кеты *Oncorhynchus keta* Walb. — одного из тетраплоидных видов лососевых. *Генетика*, 1972, 8, 2, 67.

[5] *Алтухов Ю. П., Животовский Л. А., Садыков С. С., Абдуллаев Б. А., Калабушкин Б. А.* Эффекты модального и направленного отбора по совокупности признаков у хлопчатника *Gossypium hirsutum* L. *Докл. АН СССР*, 1976, 227, 1, 212.

[6] *Баатар Д.* Анализ причин возникновения плато селекции на примере возникновения его у каракульских овец. Автореф. канд. дис. М., Ин-т общей генетики АН СССР, 1968.

[7] *Бирдмор Д. А., Шами С. А.* Гетерозиготность и оптимальный фенотип при стабилизирующем отборе. XIV Международный генетический конгресс. Тезисы докладов, 1. М., Наука, 1978, 451.

[8] *Дьячков И. Н.* Методы и организация племенного дела в каракульском овцеводстве. Алма-атинский зоовет. ин-т, 1968.

[9] *Дьячков И. Н.* Производство черных каракульских шкурок жакетных сортов. М., Колос, 1975.

[10] *Дубинин Н. П.* Экспериментальное исследование интеграции наследственных систем в процессах эволюции популяций. *Ж. общ. Биол.*, 1948, 9, 3, 203.

[11] *Дубинин Н. П., Алтухов Ю. П., Курбатова О. Л., Сусков И. И.* Интегральная генетическая характеристика «адаптивной нормы» в популяции человека. *Докл. АН СССР*, 1976, 230, 4, 957.

[12] *Егоров Е. А.* Генетические системы белков крови овец. Ташкент, ФАН, 1973.

[13] Инструкция по бонитировке каракульских ягнят с основами племенного дела. М., Колос, 1974.

[14] *Кошевой М. А.* Селекция и условия разведения каракульских овец. Ташкент, ФАН, 1975.

[15] *Lerner M. J.* Genetic homeostasis. N.Y., Wiley, 1954.

[16] *Никольский Н. Ф.* Закладка первых научных исследований по каракулеводству. В сб. Каракулеводство, 5. Самарканд, *Тр. ВНИИ каракулеводства*, 1976, 51.

[17] *Peacock A. C., Bunting S. L., Queen K. G.* Serum protein electrophoresis in acrilamide gel. *Science*, 1965, 147, 1451.

[18] *Риш М. А., Дьячков И. Н., Егоров Е. А., Мехтиев Н. Х., Ни Г. В.* Генетико-биохимическая изменчивость каракульских овец и использование ее в селекции. В сб. Каракулеводство, 5. Самарканд,. *Тр. ВНИИ каракулеводства*, 1976, 160.

[19] *Ширинский М. А.* Научные исследования по каракулеводству в Казахстане. III Международный симпозиум по каракулеводству. Самарканд, 1975, 108.
[20] *Soulè M. E.* Heterozygosity and developmental stability: another look. *Evolution*, 1978, 33, 1, 396.
[21] *Васин Б. Н.* Основные методические установки племенной работы с каракулями. *Сов. зоотехн.*, 1939, 2-3, 121.
[22] *Васин Б. Н., Васина-Попова Е. Т., Грабовский И. Н., Крымская Э. К., Петров В. А.* Руководство по каракулеводству. М., Колос, 1971.
[23] *Waddington C. H.* New patterns in genetics and development. New York, 1962.

MESURES DE LA VARIABILITÉ DES POPULATIONS SELON LES CARACTÈRES POLYMORPHES

L. Jivotovski

La variabilité dans la population et entre les populations est l'objet de nombreuses recherches dont la plupart se rattachent au *polymorphisme*: présence dans un groupe d'individus de plusieurs variantes du caractère. Nous appellerons ultérieurement ces variantes les « *morphes* » comprenant sous ce terme les allèles, les génotypes, les phénotypes, les rangs des caractères quantitatifs. Les fréquences des « morphes » sont les données de départ pour l'analyse des populations.

Les questions habituelles dans l'analyse des populations sont les suivantes: à quel point une population donnée est-elle diverse d'après les caractères polymorphes, à quel degré la distribution des fréquences des « morphes » varie-t-elle dans l'aire d'habitation de la population ou change-t-elle dans le temps, à quel point les populations diffèrent-elles, etc. Pour répondre à ces questions, on se sert de caractéristiques numériques (ou d'indices) résumées par des formules mathématiques. D'une manière générale, il existe beaucoup de formules de ce genre pour la même caractéristique (pour la diversité dans la population ou la similitude des populations par exemple). Dans certains cas, ces formules fournissent des résultats analogues et, à première vue, se distinguent peu les unes des autres. Mais elles donnent parfois des résultats tout à fait différents et on se demande alors de quelle formule faut-il se servir.

A partir de positions théoriques générales, cette question se rattache au problème de la justification des caractéristiques utilisées de la variabilité des populations. La plupart d'entre elles sont souvent justifiées par des considérations purement formelles: par « analogie ». Par exemple, la distance entre les populations est parfois fixée à partir d'analogies géométriques: chaque population est représentée sous forme de point dans l'espace multidimensionnel des « morphes », puis on calcule sur une droite la distance entre elles comme en géométrie analytique. On attribue parfois un poids différent aux axes de coordonnées ou l'on introduit une liaison entre eux et la distance

© Издательство Наука, 1979.

est alors déterminée dans un espace courbe. Lorsqu'on évalue la variété dans la population, on utilise souvent la mesure informationnelle de la diversité par analogie avec la théorie de l'information. Il est clair qu'une telle approche formelle, qui donne souvent des estimations vraisemblables, peut induire en erreur du fait que pour les différentes tâches de la biologie des populations il n'existe pas d'indice qui soit bon dans toutes les situations.

En effet, dans les recherches actuelles au niveau des populations on se livre à des comparaisons de différents types, suivant des classes différentes de caractères, avec des buts divers. On compare des groupes d'individus à l'intérieur des populations, entre les populations d'une même espèce, entre les espèces. On prend comme caractères la structure primaire des protéines, les allèles mis en évidence par des méthodes biochimiques, les caractères morphologiques qualitatifs, etc. On étudie l'influence de la mutation, de la dérive génétique, des migrations, de la sélection sur la dynamique de ces groupes. Naturellement, dans certains cas, un indice peut être meilleur ; dans d'autres cas, un second. Par exemple, dans les questions se rattachant à l'étude de la dynamique des populations sous l'influence de la mutation seule et de processus fortuits, la mesure naturelle de la diversité allélique est ce qu'on appelle nombre effectif d'allèles [2] et la mesure de la proximité des populations, la distance génétique déterminée par les fréquences des allèles [1, 3].

Dans le cas où les populations sont étudiées d'après les fréquences des phénotypes (ou même des allèles mais compte tenu de l'effet sélectif), l'emploi des indices mentionnés est purement formel, par « analogie ». Dans de telles situations, il faut avoir d'autres mesures de diversité. Comment définir ces mesures ?

Pour les besoins de la biologie des populations, nous proposons une approche méthodiquement nouvelle de l'obtention d'indices de la variabilité des populations que nous résumons ainsi : *l'indice doit satisfaire à des conditions imposées d'avance et interprétées au niveau des populations. Ces conditions doivent être traduites en langage mathématique, c'est-à-dire exprimées dans certains rapports mathématiques. Une rigoureuse analyse mathématique des rapports obtenus fera ressortir la formule satisfaisant aux conditions a priori.*

Ce principe, qui sera concrétisé plus loin pour les indices de la diversité dans la population et de la similitude des populations, souligne que l'indice ne doit pas être choisi en se basant sur des analogies de forme quelles qu'elles soient. Les données initiales seront ici les conditions interprétées au niveau des populations. L'analyse mathématique de ces conditions conduit à une formule requise pour les indices recherchés. L'important, c'est qu'ici la critique de l'indice obtenu passe à un niveau constructif : à la discussion de la justification biologique des conditions imposées et à l'« exactitude » de leur traduction en langage mathématique.

Pour l'analyse de la variabilité des populations d'après les caractères polymorphes, on peut postuler les exigences suivantes.

1° De petites variations de fréquences des morphes répandus ainsi que l'apparition ou la disparition de morphes rares ne modifient que peu la valeur de l'indice.

2° L'indice ne dépend pas de paramètres différenciant les divers morphes (l'adaptabilité des morphes, leur valeur économique, etc.).

3° La contribution des morphes à la valeur de l'indice doit être proportionnelle à leurs fréquences dans les populations.

4° L'indice doit avoir une interprétation « simple » au niveau des populations et posséder de « bonnes propriétes » statistiques.

Remarquons que cette dernière exigence est difficile à formuler: l'interprétation « simple » dont il est question n'est pas facile à traduire en termes mathématiques et il y a beaucoup de subjectivité dans l'appréciation de la meilleure et la plus simple des interprétations. Par « bonnes propriétés » statistiques on entend, en particulier, la simplicité des formules tant pour l'indice même que pour son erreur d'échantillonnage, ainsi que le « pouvoir discriminateur » de l'indice. Remarquons que la plupart des mesures de la variabilité des populations sont souvent calculées sans erreurs statistiques (étant donné qu'elles sont inconnues ou complexes), et ceci exclut la possibilité d'une comparaison juste. Les indices trouvés en raison des exigences énumérées se sont avérés simples par leur structure et pour eux on peut calculer l'erreur d'échantillonnage. Il importe aussi de noter que les indices obtenus sont généralisés pour le cas des caractères quantitatifs, ce qui permet d'analyser dans les populations les caractères qualitatifs et quantitatifs à l'aide d'une seule méthode.

Nous présentons ci-dessous les indices permettant d'évaluer la diversité dans la population et la similitude des populations d'après les caractères polymorphes et qui font face aux exigences 1°-4°. Les formules mathématiques résumant ces exigences et l'analyse mathématique correspondante ont été exposées par l'auteur de cet article dans d'autres ouvrages [4, 5].

L'indice de diversité dans la population

Supposons que m « morphes » aient été découverts dans une population. Désignons par $p_1, p_2, \ldots, p_m$ les valeurs échantillonnées des fréquences ($p_1 + p_2 + \ldots + p_m = 1$), N étant l'effectif de l'échantillon. La diversité dans la population peut être estimée par l'indice μ appelé *nombre moyen de morphes* dans la population :

$$\mu = (\sqrt{p_1} + \sqrt{p_2} + \ldots + \sqrt{p_m})^2.$$

L'indice μ donne une estimation de la diversité en unités « nombre de morphes ». Sa valeur maximale possible est égale à m quand la fréquence de tous les morphes est la même (c'est-à-dire quand $p_1 =$

$= 1/m$, $p_2 = 1/m$, . . ., $p_m = 1/m$). Lorsque la distribution des fréquences des morphes est inégale, $\mu < m$. En cas de monomophisme, $\mu = 1$. La formule approchée pour le calcul de l'erreur statistique (d'échantillonnage) s_μ est:

$$s_\mu \approx \sqrt{\frac{\mu(m-\mu)}{N}}.$$

Si l'on étudie n caractères indépendants, on peut définir le nombre moyen de morphes dans la population d'après l'ensemble de ces caractères comme une moyenne arithmétique:

$$\bar{\mu} = \frac{1}{n}(\mu_1 + \ldots + \mu_n), \quad s_\mu \approx \frac{1}{n}\sqrt{s_1^2 + \ldots + s_n^2}$$

où μ_1, μ_2, ... μ_n et s_1, s_2, . . ., s_n sont respectivement les valeurs de l'indice de diversité et les erreurs pour différents caractères.

La part des morphes rares dans la population

A côté du nombre moyen de morphes, on peut définir l'indice h, part des morphes rares:

$$h = 1 - \mu/m, \quad s_h \approx \sqrt{\frac{h(1-h)}{N}}.$$

L'indice h donne une information nouvelle, par comparaison à μ, sur le caractère de la diversité dans la population. Alors que μ montre le degré de diversité de la population, l'indice h évalue la structure de cette diversité.

Exemple. Supposons que pour les échantillons tirés de trois populations (hypothétiques) on ait obtenu les fréquences des phénotypes présentées au *Tableau 1.*

Tableau 1

Caractéristiques de la diversité dans trois populations hypothétiques

Population	Fréquences des phénotypes					Effectif, N	Nombre moyen de phénotypes, μ	Part des phénotypes rares, h
	p_1	p_2	p_3	p_4	p_5			
I	0,45	0,20	0,20	0,15		100	3,81±0,085	0,047±0,021
II	0,60	0,20	0,14	0,04	0,02	50	3,75±0,306	0,250±0,061
III	0,30	0,15	0,25	0,20	0,10	60	4,83±0,117	0,034±0,023

Ici la différence entre les populations I et II d'après le nombre moyen de phénotypes est non significative, c'est-à-dire que la degré de diversité dans ces populations est le même. Cependant, le caractè-

re de la diversité diffère: dans la population II la part des phénotypes rares est élevée, la population I est plus homogène à l'égard des fréquences des phénotypes. Dans la population III, les phénotypes rares sont presque absents, mais la diversité y est plus élevée que dans la population I.

Si l'on étudie n caractères indépendants statistiquement, on peut déterminer la part des morphes rares pour l'ensemble de ces caractères:

$$\overline{h} = 1 - \overline{\mu}/\overline{m}, \quad s_{\overline{h}} \approx s_{\overline{\mu}}/\overline{m}$$

où $\overline{\mu}$ et s_{μ} sont définis plus haut,

$\overline{m} = {}^1/_n (m_1 + m_2 + \ldots + m_n)$,

$m_1, m_2, \ldots, m_n$ est le nombre de morphes découverts pour les caractères correspondants.

L'indice de similitude des populations

Supposons que sont comparées deux populations dans lesquelles on a découvert m morphes différents. Leur fréquence dans l'échantillon de la première population est désignée par $p_1, p_2, \ldots, p_m$ et dans l'échantillon de la seconde population par $q_1, q_2, \ldots, q_m$. Les effectifs des échantillons sont respectivement N_1 et N_2. L'indice de similitude des populations r est trouvé selon la formule

$$r = \sqrt{p_1 q_1} + \sqrt{p_2 q_2} + \ldots + \sqrt{p_m q_m}.$$

La valeur de l'indice de similitude ne dépasse pas 1. Elle est égale à 1 dans le cas où les populations comparées sont identiques (par les fréquences des morphes). L'indice r est égal à zéro quand les populations comparées ne possèdent pas un seul morphe commun. Dans les autres cas, ses valeurs sont comprises entre 0 et 1.

L'indice de similitude r est la mesure de la similitude des populations comparées par paires. On peut l'interpréter comme la fréquence de morphes communs dans les populations comparées, étant donné qu'un apport non nul à la valeur de r n'est fait que par les morphes communs aux deux populations.

Passons maintenant à l'évaluation de r. Etant donné que dans les échantillons comparés, les fréquences des morphes correspondants peuvent être différentes, ne serait-ce que parce que les échantillons sont tirés au hasard, presque toujours $r < 1$ (même si les échantillons sont prélevés sur un même ensemble général). Aussi peut-il d'abord décider si r diffère significativement de 1, autrement dit si les populations comparées sont différentes. D'une façon générale, il est facile de répondre à cette question en appliquant un des tests connus de comparaison des échantillons (test de χ^2 par exemple). Cependant, cette approche recèle une certaine incommodité: si les échantillons diffèrent significativement les uns des autres et si la question de l'estima-

tion de leur degré de similitude devient importante, il faudra, outre le calcul de χ^2, chercher la valeur de r et son erreur d'échantillonnage. Afin d'éviter un double travail, indiquons un nouveau test de comparaison des échantillons : *critère d'identité I* ayant pour base l'indice de similitude r:

$$I = \frac{8N_1N_2}{N_1+N_2}\left(1-r-\frac{n_1+n_2}{2(N_1+N_2)}\right),$$

où $n_1 = p^0N_1$, $n_2 = q^0N_2$,

p^0 est la somme des fréquences des morphes du 1er échantillon absents dans le 2e échantillon,

q^0 est la somme des fréquences des morphes du 2e échantillon absents dans le 1er.

Dans le cas où chaque morphe est présent dans les deux échantillons, on aura $p^0 = 0$, $q^0 = 0$ et le critère d'identité se simplifie :

$$I = \frac{8N_1N_2}{N_1+N_2}(1-r).$$

La grandeur I est distribuée à peu près comme χ^2 avec $m - 1$ degrés de liberté dans l'hypothèse nulle où les deux échantillons sont tirés d'un même ensemble général. Si I dépasse la valeur de table de χ^2 au niveau de signification fixé, il existe une différence entre les échantillons (au niveau de signification correspondant).

Si r diffère significativement de 1 (d'après le critère d'identité I), il faut calculer son erreur d'échantillonnage selon la formule

$$s_r \approx \frac{1}{2}\sqrt{\frac{1-q^0-r^2}{N_1}+\frac{1-p^0-r^2}{N_2}}.$$

Dans le cas où chaque morphe est présent dans les deux échantillons, c'est-à-dire que $p^0 = 0$, $q^0 = 0$, la formule se simplifie :

$$s_r \approx \frac{1}{2}\sqrt{\frac{N_1+N_2}{N_1N_2}(1-r^2)}.$$

Exemple. Pour les caractéristiques données au *Tableau 1* les valeurs de l'indice de similitude r et du critère d'identité I sont présentées au *Tableau 2*.

Le *Tableau 2* montre que les échantillons I et II diffèrent l'un de l'autre et que l'échantillon III est significativement différent des deux premiers.

Si l'on étudie n caractères indépendants statistiquement, on peut définir la similitude moyenne des populations d'après l'ensemble des caractères comme une moyenne arithmétique :

$$\bar{r} = \frac{1}{n}(r_1+r_2+\ldots+r_n), \quad s_{\bar{r}} \approx \frac{1}{n}\sqrt{s_1^2+s_2^2+\ldots+s_n^2}.$$

où $r_1, r_2, \ldots, r_n$ et $s_1, s_2, \ldots, s_n$ sont les valeurs correspondantes de l'indice de similitude et les erreurs pour différents caractères.

Tableau 2

Indice de similitude et critère d'identité pour les populations du *Tableau 1* (les valeurs de table de χ^2 à 4 degrés de liberté sont égales à 9,49 au niveau de signification 5 % et à 13,28 au niveau de signification 1 %)

Echantillons comparés	I-II	I-III	II-III
I	8,8	13,2	17,7
$r \pm s_r$	0,964	0,937±0,0237	0,919±0,0377

Le critère d'identité pour $\bar{r}$ est égal à la somme des valeurs du critère pour différents caractères: $I = I_1 + I_2 + \ldots + I_n$ et se vérifie pour la signification d'après la table de χ^2 avec le nombre total de degrés de liberté $m = m_1 + m_2 + \ldots + m_n - n$ où m_1, m_2, $\ldots$ m_n sont les nombres de morphes présents dans les échantillons comparés pour différents caractères.

Indice de similitude de plusieurs populations

Dans le cas où la comparaison s'étend à un grand nombre d'échantillons, on peut estimer, en s'appuyant sur l'indice r, toutes les similitudes possibles des échantillons appariés et les représenter graphiquement, par exemple, sous la forme d'un dendrogramme. Cependant, dans plusieurs cas, il est important d'obtenir, outre toutes les similitudes possibles des échantillons appariés, un indice unique reflétant la similitude de l'ensemble d'échantillons entre eux. Pour trouver cet indice unique compte tenu de r, on peut proposer l'approche suivante. On commence par faire la moyenne des fréquences des morphes sur tous les k échantillons. Puis, on introduit un échantillon conventionnel qui correspond à ces fréquences « moyennées ». On l'appelle « échantillon moyenné ». Pour chacun des k échantillons on cherche la similitude r par rapport à ce dernier. Enfin, on fait la moyenne sur toutes les k valeurs trouvées de r. L'indice ainsi obtenu évalue, par exemple, la similitude moyenne des populations étudiées par rapport à une certaine population mère — le point de divergence (compte tenu des restrictions et des suppositions). L'approche décrite peut être utilisée pour la comparaison de populations d'une même espèce « non liées » ou de populations d'espèces différentes.

Dans d'autres cas, lors de l'étude de la structure au sein de la population par exemple, une question peut surgir, celle de savoir à quel point la distribution des fréquences du caractère polymorphe varie dans le temps ou dans l'espace (dans l'aire d'habitation par

exemple). Dans ce cas, on introduit l'indice de similitude R qui est un r généralisé pour le cas de k échantillons. Son interprétation est la même que pour r: l'indice R estime la fréquence de morphes communs pour tous les k échantillons. Voici comment on le calcule.

Supposons que $p_{11}, p_{12}, \ldots, p_{1m}$ soient les fréquences des morphes dans le 1er échantillon, $p_{21}, p_{22}, \ldots, p_{2m}$ les fréquences dans le 2e échantillon, etc., $p_{k1}, p_{k2}, \ldots, p_{km}$ les fréquences des morphes dans le ke échantillon. L'indice de similitude générale R est trouvé d'après la formule

$$R = \sqrt[k]{p_{11}p_{21}\ldots p_{k1}} + \sqrt[k]{p_{12}p_{22}\ldots p_{k2}} + \ldots + \sqrt[k]{p_{1m}p_{2m}\ldots p_{km}}.$$

La formule approchée pour l'erreur d'échantillonnage est:

$$s_R \approx \sqrt{\frac{1-R^2}{k\tilde{N}}}$$

où $\tilde{N}$ est la moyenne harmonique des effectifs des échantillons comparés:

$$1/\tilde{N} = \frac{1}{k}(1/N_1 + 1/N_2 + \ldots + 1/N_k).$$

Il y a lieu de noter que lorsque le nombre d'échantillons est grand et que leurs effectifs sont relativement petits, la valeur de R peut avoir un écart significatif: elle est moindre devant la valeur générale théorique. On donne ci-après la formule approchée pour la correction ΔR:

$$\Delta R \approx \frac{(m-R)(k-1)}{2kN}.$$

En définitive, la valeur de l'indice de similitude générale sera $R + \Delta R$. Les formules précises de s_R et ΔR ne sont pas connues.

Si l'on étudie n caractères indépendants, on peut définir la similitude générale moyenne d'après l'ensemble des caractères comme une moyenne arithmétique:

$$\bar{R} = \frac{1}{n}(R_1 + \ldots + R_n), \quad s_{\bar{R}} \approx \sqrt{\frac{1}{n}(s_1^2 + \ldots + s_n^2)}.$$

RÉFÉRENCES

[1] *Cavalli-Sforza L. L., Bodmer W. F.* The Genetics of human populations. San Francisco, W. H. Freeman and Co., 1971.

[2] *Crow J. F., Kimura M.* An introduction to population genetics theory. N.Y., Harper and Row, 1970.

[3] *Nei M.* Molecular population genetics and evolution. Amsterdam, 1975.

[4] *Животовский Л. А.* Показатель сходства популяций по полиморфным признакам. *Журн. общ. Биол.*, 1979, 40, 4, 587-602.

[5] *Животовский Л. А.* Показатель внутрипопуляционного разнообразия. *Журн. общ. Биол.*, 1980, 41, 6, 828-836.

MESURES DE LA VARIABILITÉ DES POPULATIONS SELON UN ENSEMBLE DE CARACTÈRES QUANTITATIFS

L. Jivotovski

En génétique des populations, l'analyse d'un ensemble de caractères quantitatifs est embarrassée par la présence de corrélations. Généralement, l'action de n'importe quel facteur porte non pas sur les caractères pris en particulier, mais sur toute leur totalité. Aussi les conclusions tirées de l'étude d'un seul caractère sont correlées à celles déduites pour tous les autres caractères. Ceci peut conduire à des jugements erronés. Pour l'éviter, on recourt souvent à la substitution de nouveaux caractères non correlés aux caractères initiaux, par exemple, des indices résumant les proportions des dimensions ou des combinaisons linéaires de caractères (des composantes principales). Cette substitution, importante par elle-même, ne lève pourtant pas le problème, car on juge de l'absence de corrélations entre les caractères d'après une classe d'échantillons déterminée et ce jugement peut être incorrect pour une autre classe d'observations. Par exemple, des caractères indépendants au sein d'une subpopulation peuvent être connexes à l'échelon d'hiérarchie suivant, quand l'unité d'observation est déjà la subpopulation. Des corrélations peuvent apparaître, se modifier ou disparaître dans le cas du groupe d'âge suivant; des corrélations génétiques peuvent manquer en présence de corrélations phénotypiques et *vice versa*. C'est pourquoi l'étude détaillée de chacun des nouveaux caractères pris en particulier, de même que des caractères initiaux, exige de toute façon une analyse corrélative.

Mais, lors de l'analyse d'un grand ensemble de caractères, l'information de sortie est abondante : pour 10 caractères, le nombre de toutes les corrélations possibles est égal à 45 ; pour 15, il dépasse 100. Dans le cas général de p caractères, il est égal à $\frac{p(p-1)}{2}$. Dans cette situation, l'analyse détaillée des corrélations est pratiquement impossible. Il est clair que l'étude du complexe de caractères quantitatifs est impensable sans l'introduction d'un petit nombre de paramètres intégraux. L'idée n'est pas neuve et elle a toujours surgi

© Издательство Наука, 1980.

dans l'analyse d'une totalité des caractères. Des caractéristiques telles que le coefficient de sélection, les composantes principales, les facteurs, etc., réduisent la totalité des données à un seul ou à un petit nombre d'indices. Néanmoins, cette énumération des problèmes résolus est loin d'être complète. Un des principaux problèmes — l'analyse de la variabilité entre groupes et dans les groupes d'après l'ensemble des caractères quantitatifs — n'est pas élaboré au niveau de la génétique des populations bien que, pour le cas d'un seul caractère, cette analyse soit depuis longtemps traditionnelle.

Dans l'article présent, on introduit des paramètres généralisés pour un ensemble de caractères quantitatifs : variance généralisée, indice généralisé d'intensité de l'influence du facteur, coefficient de corrélation généralisé de tous les caractères d'un ensemble donné, indice de similitude généralisé des populations. Les paramètres généralisés introduits définissent le complexe de caractères analysé comme un entier. Ici, la présence de corrélations entre les caractères n'est pas gênante, au contraire, c'est une part importante de l'information sur la variabilité d'un complexe de caractères donné.

Nous utiliserons ultérieurement les désignations suivantes: p est le nombre de caractères; le vecteur de caractères x constitué de p composantes est l'unité de mesure déterminant l'individu (ou tout autre objet) de façon que $x = (x_1, x_2, \ldots, x_p)$ où x_1 est la valeur du 1er caractère, x_2 la valeur du 2e caractère, etc., x_p la valeur du p-ième caractère d'un individu donné. Ordinairement, x est interprété géométriquement comme un point dans l'espace p-dimensionnel des caractères. Pour un échantillon d'effectif N la totalité des observations (vecteurs) est comprise dans un certain domaine de cet espace. Le « centre » et la « diamètre » de ce domaine caractérisent respectivement la valeur moyenne des caractères et leur variabilité pour un échantillon donné. Pour un autre échantillon, la position du domaine dans l'espace des caractères peut être différente. Son centre ainsi que son diamètre peuvent aussi différer. Dans cette interprétation, les questions de l'analyse de la variabilité de la totalité des caractères sont formulées comme les problèmes de l'analyse de la position réciproque des domaines caractérisant les échantillons dans l'espace des caractères. Les méthodes d'étude de l'ensemble des caractères reposent donc sur les principes de l'analyse statistique multidimensionnelle. Signalons qu'une telle approche est liée à des opérations sur des matrices et des vecteurs et exige donc l'emploi de la technique de calcul moderne.

Les matrices de covariances et de corrélations

Les notions de départ d'une analyse statistique multidimensionnelle sont celles de matrices de covariances et de corrélations. Rappelons-les. Soit x_i le vecteur des caractères d'un individu i: $x_i =$

$= (x_{i1}, x_{i2}, \ldots, x_{ip})$ où x_{ij} est la valeur du caractère j de l'individu i ($j = 1, 2, \ldots, p$). Appelons N l'effectif ou la taille de l'échantillon. Soit $N > p$.

Désignons par $\bar{x}_j = \frac{1}{N} \sum_i x_{ij}$ la valeur moyenne du caractère j dans l'échantillon. La variance de ce caractère est égale à

$$\sigma_j^2 = \frac{1}{N-1} \left(\sum_i x_{ij}^2 - N\bar{x}_j^2 \right)$$

Introduisons la covariance s_{jl} des caractères j et l:

$$s_{jl} = \frac{1}{N-1} \left(\sum_i x_{ij} x_{il} - N\bar{x}_j \bar{x}_l \right).$$

La covariance est intimement liée à la notion de coefficient de corrélation paire r_{jl} entre les caractères j et l:

$$r_{jl} = \frac{s_{jl}}{\sigma_j \sigma_l}.$$

De la définition de la variance et de la covariance on déduit que la variance du caractère j est une covariance de ce caractère avec lui-même. Aussi, pour unifier les notations, il sera commode, par la suite, d'utiliser parfois s_{jj} au lieu de σ_j^2.

A partir de toutes les covariances et variances possibles des caractères on peut établir la matrice de covariances

$$S = \begin{Vmatrix} s_{11} & s_{12} & \ldots & s_{1p} \\ s_{21} & s_{22} & \ldots & s_{2p} \\ \ldots & & & \\ s_{p1} & s_{p2} & \ldots & s_{pp} \end{Vmatrix}$$

Les variances des caractères se trouvent sur la diagonale principale de cette matrice. L'élément hors de la diagonale à l'intersection de la ligne j et de la colonne l est la covariance des caractères j et l.

Par analogie, on détermine la matrice de corrélations R:

$$R = \begin{Vmatrix} 1 & r_{12} & \ldots & r_{1p} \\ r_{21} & 1 & \ldots & r_{2p} \\ \ldots & \ldots & \ldots & \ldots \\ r_{p1} & r_{p2} & \ldots & 1 \end{Vmatrix}$$

Les unités se trouvent sur la diagonale principale de la matrice de corrélations. L'élément hors de la diagonale à l'intersection de la ligne j et de la colonne l est le coefficient de corrélation des caractères j et l. Les matrices de corrélations et de covariances sont symétriques par rapport à la diagonale principale. Les matrices de covarian-

ces et de corrélations sont reliées entre elles puisque, comme nous l'avons dit, $s_{jl} = r_{jl}\sigma_j\sigma_l$.

Par la suite, il sera commode d'utiliser la notation matricielle pour déterminer la matrice de covariances. Désignons par x^T le vecteur colonne:

$$x^T = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}.$$

Partant des règles de la sommation des vecteurs, nous obtenons $\bar{x} = \frac{1}{N}\sum_i x_i$ où x_i est le vecteur des caractères de l'individu i. Partant des règles de la multiplication des vecteurs et des matrices, nous avons:

$$S = \frac{1}{N-1}\left(\sum_i x_i^T x_i - N\bar{x}^T\bar{x}\right).$$

Il y a une analogie évidente de la notation avec la définition de la variance d'un caractère. L'exposé détaillé de la théorie des matrices appliquée à l'analyse statistique multidimensionnelle et des notions des matrices de covariances et de corrélations est donné dans plusieurs monographies [1, 4, 5, 6].

A titre d'exemple, examinons deux caractères, soit $p = 2$. Alors r_{12} et r_{21} sont les coefficients de corrélation entre les caractères. Désignons-les simplement par r. Alors $R = \begin{Vmatrix} 1 & r \\ r & 1 \end{Vmatrix}$. Etant donné que $s_{11} = \sigma_1^2$, $s_{22} = \sigma_2^2$, $s_{12} = s_{21} = r\sigma_1\sigma_2$, alors

$$S = \begin{Vmatrix} \sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_1\sigma_2 & \sigma_2^2 \end{Vmatrix}.$$

Nous aurons encore besoin de la notion de déterminant des matrices S et R que nous désignerons par det (S) et det (R). Son examen général est donné dans les monographies indiquées plus haut. Pour le cas où $p = 2$, det $(S) = \sigma_1^2\sigma_2^2(1 - r^2)$ et det $(R) = 1 - r^2$. On voit que les déterminants des matrices de covariances et de corrélations sont reliés entre eux par la variance des caractères: det $(S) = \sigma_1^2\sigma_2^2$ det (R). Dans le cas général de p caractères, cette connexion est conservée: det $(S) = \sigma_1^2\sigma_2^2 \ldots \sigma_p^2$ det (R).

Variance généralisée de l'ensemble des caractères

Beaucoup d'auteurs (des monographies citées par exemple) proposent de prendre comme caractéristique de variabilité de la totalité des caractères le déterminant de la matrice de covariances:

det (S). A notre avis, cela ne correspond pas toujours aux buts des recherches étant donné que les paramètres estimés ainsi ne peuvent être comparés entre eux s'ils sont obtenus pour les ensembles ayant un nombre de caractères différent. Nous introduisons en l'occurrence une condition.

Exigence. *La variance généralisée de l'ensemble des caractères doit être normalisée par rapport au nombre de caractères.*

Pour satisfaire cette exigence, définissons la valeur générale de la variance généralisée $\hat{\sigma}^2$ comme suit:

$$\hat{\sigma}^2 = \sqrt[p]{\det(\hat{S})}.$$

Ici et dans la suite le signe « ˆ » signifiera la valeur générale du paramètre.

Soit $p = 2$. Alors

$$\hat{\sigma}^2 = \sqrt{\det(\hat{S})} = \sqrt{\hat{\sigma}_1^2\hat{\sigma}_2^2(1-\hat{r}^2)} = \hat{\sigma}_1\hat{\sigma}_1\sqrt{1-\hat{r}^2}.$$

Admettons que les variances des deux caractères soient identiques: $\hat{\sigma}_1^2 = \hat{\sigma}_2^2$. Alors $\hat{\sigma}^2 = \hat{\sigma}_1^2\sqrt{1-\hat{r}^2}$. Par conséquent, la variance généralisée dépend de la corrélation entre les caractères et l'on a toujours $\hat{\sigma}^2 < \hat{\sigma}_1^2$. Si les caractères sont indépendants, la variance généralisée coïncide avec la variance ordinaire du caractère: $\hat{\sigma}^2 = \hat{\sigma}_1^2$. Dans le cas de variances inégales pour $\hat{r} = 0$, la variance généralisée est égale à la moyenne géométrique des variances des caractères initiaux: $\hat{\sigma}^2 = \sqrt{\hat{\sigma}_1^2\hat{\sigma}_2^2} = \hat{\sigma}_1\hat{\sigma}_2$. De façon analogue, dans le cas général où tous les p caractères sont égaux : $\hat{\sigma}^2 = \sqrt[p]{\hat{\sigma}_1^2\hat{\sigma}_2^2 \ldots \hat{\sigma}_p^2}$, c'est-à-dire que la variance généralisée est égale à la moyenne géométrique des variances des caractères initiaux. S'il y a une corrélation entre les caractères, alors

$$\hat{\sigma}^2 = \sqrt[p]{\hat{\sigma}_1^2\hat{\sigma}_2^2 \ldots \hat{\sigma}_p^2} \cdot \sqrt[p]{\det(\hat{R})}.$$

Du fait que toujours $\det(\hat{R}) < 1$, on peut déduire que la présence d'une corrélation fait diminuer la variance généralisée. C'est tout naturel, car les caractères de l'ensemble varient non pas indépendamment les uns des autres, mais dans les limites des liaisons qui bornent leur variation.

Estimation statistique de la variance généralisée

Comme nous l'avons vu plus haut, la variance généralisée peut être calculée d'après la matrice de covariances comme $\sqrt[p]{\det(S)}$ (ou d'après la matrice de corrélations comme $\sqrt[p]{\sigma_1^2 \ldots \sigma_p^2} \cdot \sqrt[p]{\det(R)}$). Cependant, la matrice de covariances est pratiquement l'estimation d'échantillonnage de sa valeur générale. Aussi $\sqrt[p]{\det(S)}$ est-elle une estimation biaisée de la variance généralisée. La valeur de l'écart dépend de l'effectif de l'échantillon et tend vers zéro quand N est beaucoup plus grand que p. Pour de petits effectifs N, cet écart peut être éliminé en utilisant la grandeur G figurant dans le *Tableau 1*:

$$\sigma^2 = G \sqrt[p]{\det(S)} \tag{1}$$

où G est la valeur de table correspondant au nombre de caractères p et à l'effectif de l'échantillon N.

Dans la suite, nous admettrons que G a pour paramètres p et N et nous la désignerons pour plus de brièveté par $G_{p;N}$.

Soit, par exemple, $p = 7$, $N = 12$ et soit $\sqrt[7]{\det(S)} = 6{,}7$. D'après le *Tableau 1*, $G = 1{,}6$, aussi la variance généralisée est-elle égale à $\sigma^2 = 10{,}72$.

Le *Tableau 1* est dressé en se basant sur une formule exacte (21) (v. *Annexe mathématique*). Si N dépasse la valeur de table, on peut se servir de cette formule. Cependant, elle n'est pas commode à l'usage et nous donnerons la formule approchée suivante pour le calcul de G:

$$G = \frac{N}{N-1} \exp\left\{\frac{p}{2N} + \frac{p^2}{6N^2}\right\} \quad (N \gg p). \tag{2}$$

L'*Annexe mathématique* montre que la grandeur $\nu\sigma^2/\hat{\sigma}^2$ est distribuée à peu près comme χ^2 avec ν degrés de liberté. Le nombre de degrés de liberté est figuré dans le *Tableau 1*. Si N dépassent les valeurs de table, on peut utiliser la formule approchée

$$\nu = \frac{p^2}{\ln\left(\frac{N-1}{N-p-1}\right)} \quad (N \gg p). \tag{3}$$

Etant donné que la grandeur $\nu\sigma^2/\hat{\sigma}^2$ est distribuée comme χ^2_ν, les méthodes statistiques pour la variance généralisée sont identiques aux méthodes dont on dispose pour un seul caractère. L'erreur statistique de la variance (m_{σ^2}) et l'erreur de l'écart type (m_σ) sont égales:

$$m_{\sigma^2} = \sqrt{\frac{2}{\nu}}\,\hat{\sigma}^2 \text{ et } m_\sigma = \sqrt{\frac{1}{2\nu}}\,\hat{\sigma}.$$

Tableau 1

Corrections *G* et *degrés de liberté* ν pour la variance généralisée (les valeurs de *N* sont portées de $p+1$ à $p+30$)

N	$p=2$		$p=3$		$p=4$		$p=5$		$p=6$		$p=7$		$p=8$	
	G	ν	*G*	ν	*G*	ν	*G*	ν	*G*	ν	*G*	ν	*G*	ν
$p+1$	2,00	2,00	2,48	3,43	2,72	5,22	2,83	7,34	2,89	9,79	2,92	12,54	2,94	15,60
$p+2$	1,50	4,00	1,77	6,61	1,94	9,78	2,05	13,46	2,14	17,64	2,20	22,29	2,25	27,41
$p+3$	1,33	6,00	1,52	9,71	1,66	14,07	1,75	19,06	1,83	24,63	1,89	30,77	1,94	37,46
$p+4$	1,25	8,00	1,40	12,77	1,51	18,26	1,59	24,43	1,66	31,25	1,72	38,70	1,77	46,75
$p+5$	1,20	10,00	1,32	15,81	1,41	22,38	1,48	29,68	1,55	37,67	1,60	46,34	1,65	55,65
$p+6$	1,17	12,00	1,27	18,84	1,35	26,47	1,41	34,86	1,47	43,98	1,52	53,81	1,56	64,32
$p+7$	1,14	14,00	1,23	21,86	1,30	30,53	1,36	39,99	1,41	50,21	1,45	61,17	1,50	72,83
$p+8$	1,12	16,00	1,20	24,87	1,26	34,58	1,32	45,10	1,36	56,40	1,41	68,45	1,44	81,24
$p+9$	1,11	18,00	1,18	27,89	1,24	38,63	1,28	50,19	1,33	62,55	1,37	75,69	1,40	89,58
$p+10$	1,10	20,00	1,16	30,90	1,21	42,66	1,26	55,26	1,30	68,68	1,33	82,88	1,37	97,85
$p+11$	1,09	22,00	1,15	33,91	1,20	46,69	1,24	60,32	1,27	74,78	1,31	90,04	1,34	106,1
$p+12$	1,08	24,00	1,14	36,91	1,18	50,71	1,22	65,37	1,25	80,37	1,28	97,18	1,31	114,3
$p+13$	1,08	26,00	1,13	39,92	1,17	54,73	1,20	70,42	1,23	86,95	1,26	104,3	1,29	122,5
$p+14$	1,07	28,00	1,12	42,92	1,16	58,75	1,19	75,45	1,22	93,01	1,25	111,4	1,27	130,6
$p+15$	1,07	30,00	1,11	45,93	1,15	62,77	1,18	80,49	1,21	99,07	1,23	118,5	1,26	138,8
$p+16$	1,062	32,00	1,103	48,933	1,136	66,780	1,166	85,517	1,194	105,12	1,219	125,58	1,244	146,882
$p+17$	1,059	34,00	1,097	51,937	1,128	70,792	1,157	90,543	1,183	111,17	1,208	132,66	1,231	154,991
$p+18$	1,056	36,00	1,092	54,941	1,121	74,803	1,148	95,567	1,173	117,21	1,197	139,72	1,219	163,089
$p+19$	1,053	38,00	1,087	57,944	1,115	78,813	1,141	100,58	1,165	123,25	1,187	146,78	1,209	171,178
$p+20$	1,050	40,00	1,082	60,946	1,110	82,822	1,134	105,60	1,157	129,28	1,179	153,84	1,199	179,259
$p+21$	1,048	42,00	1,079	63,949	1,105	86,830	1,128	110,62	1,150	135,31	1,171	160,89	1,190	187,333
$p+22$	1,045	44,00	1,075	66,951	1,100	90,837	1,122	115,64	1,143	141,34	1,163	167,93	1,182	195,401
$p+23$	1,043	46,00	1,072	69,953	1,096	94,844	1,117	120,65	1,138	147,37	1,157	174,98	1,175	209,464
$p+24$	1,042	48,00	1,069	72,955	1,092	98,850	1,113	125,66	1,132	153,39	1,151	182,01	1,168	211,522
$p+25$	1,040	50,00	1,066	75,957	1,088	102,86	1,108	130,68	1,127	159,41	1,145	189,05	1,162	219,576
$p+26$	1,038	52,00	1,064	78,958	1,085	106,86	1,04	135,69	1,122	165,44	1,140	196,08	1,156	227,626
$p+27$	1,037	54,00	1,061	81,960	1,082	110,87	1,100	140,70	1,118	171,46	1,135	203,12	1,151	235,672
$p+28$	1,036	56,00	1,059	84,961	1,079	114,87	1,097	145,71	1,114	177,47	1,130	210,14	1,146	243,716
$p+29$	1,034	58,00	1,057	87,963	1,076	118,87	1,094	150,72	1,110	183,49	1,126	217,17	1,141	251,756
$p+30$	1,033	60,00	1,055	90,964	1,074	122,88	1,091	155,73	1,107	189,51	1,122	224,20	1,137	259,794

Suite

N	$p=9$		$p=10$		$p=11$		$p=12$		$p=13$		$p=14$		$p=15$	
	G	v	G	v	G	v	G	v	G	v	G	v	G	v
$p+1$	2,95	18,95	2,95	22,59	2,95	26,52	2,95	30,73	2,94	35,22	2,94	39,98	2,93	45,02
$p+2$	2,95	32,98	2,32	38,99	2,35	45,43	2,37	52,29	2,39	59,58	2,41	67,27	2,42	75,37
$p+3$	1,99	44,68	2,03	52,43	2,06	60,69	2,09	69,44	2,12	78,69	2,14	88,43	2,16	98,64
$p+4$	1,81	55,40	1,85	64,63	1,88	74,41	1,92	84,76	1,94	95,64	1,97	107,1	2,00	119,0
$p+5$	1,69	65,60	1,73	76,17	1,76	87,34	1,79	99,10	1,82	111,4	1,85	124,4	1,87	137,8
$p+6$	1,60	75,49	1,64	87,32	1,67	99,79	1,70	112,9	1,73	126,6	1,76	140,9	1,78	155,8
$p+7$	1,53	85,19	1,57	98,22	1,60	111,9	1,63	126,3	1,66	141,3	1,68	156,9	1,71	173,1
$p+8$	1,48	94,74	1,51	108,9	1,54	123,8	1,57	139,4	1,60	155,6	1,62	172,5	1,65	190,6
$p+9$	1,44	104,2	1,47	119,5	1,50	135,6	1,52	152,3	1,55	169,8	1,57	187,8	1,59	206,6
$p+10$	1,40	113,6	1,43	130,0	1,46	147,2	1,48	165,1	1,51	183,7	1,53	203,0	1,55	223,0
$p+11$	1,37	122,9	1,40	140,5	1,42	158,8	1,45	177,8	1,47	197,5	1,49	218,0	1,51	239,1
$p+12$	1,34	132,2	1,37	150,8	1,39	170,2	1,42	190,4	1,44	211,3	1,46	232,9	1,48	255,1
$p+13$	1,32	141,4	1,34	161,2	1,37	181,7	1,39	202,9	1,41	224,9	1,43	247,6	1,45	271,0
$p+14$	1,30	150,6	1,32	171,4	1,35	193,0	1,37	215,4	1,39	238,5	1,41	262,3	1,43	286,8
$p+15$	1,28	159,8	1,31	181,7	1,33	204,3	1,35	227,8	1,37	252,0	1,39	276,9	1,41	302,6
$p+16$	1,267	168,99	1,289	191,92	1,310	215,64	1,330	240,13	1,350	265,40	1,368	291,43	1,386	318,20
$p+17$	1,253	178,15	1,274	202,12	1,294	226,90	1,314	252,46	1,333	278,81	1,351	305,92	1,368	333,79
$p+18$	1,240	187,28	1,261	212,30	1,280	238,13	1,299	264,76	1,317	292,17	1,335	320,36	1,351	349,32
$p+19$	1,229	196,41	1,249	222,47	1,268	249,35	1,286	277,03	1,303	305,51	1,320	334,77	1,336	364,81
$p+20$	1,219	205,52	1,238	232,62	1,256	260,55	1,273	289,28	1,290	318,82	1,307	349,14	1,323	380,25
$p+21$	1,209	214,62	1,228	242,76	1,245	271,73	1,262	301,51	1,278	332,10	1,294	363,49	1,310	395,66
$p+22$	1,201	223,72	1,218	252,89	1,235	282,89	1,252	313,72	1,268	345,36	1,283	377,61	1,298	411,05
$p+23$	1,193	232,81	1,210	263,01	1,226	294,05	1,242	325,91	1,258	358,60	1,273	392,10	1,287	426,40
$p+24$	1,185	241,89	1,202	273,12	1,218	305,19	1,233	338,10	1,248	371,83	1,263	406,37	1,277	441,73
$p+25$	1,179	250,96	1,195	283,22	1,210	316,32	1,225	350,26	1,240	385,04	1,254	420,63	1,268	457,03
$p+26$	1,172	260,03	1,188	293,31	1,203	327,45	1,217	362,42	1,232	398,23	1,245	434,37	1,259	472,32
$p+27$	1,166	269,10	1,181	303,40	1,196	338,56	1,210	374,57	1,224	411,42	1,237	449,09	1,251	487,59
$p+28$	1,161	278,16	1,176	313,48	1,190	349,67	1,204	386,71	1,217	424,59	1,230	463,30	1,243	502,84
$p+29$	1,156	287,22	1,170	323,56	1,184	360,77	1,197	398,84	1,210	437,75	1,223	477,50	1,236	518,08
$p+30$	1,151	296,27	1,165	333,63	1,178	371,87	1,191	410,96	1,204	450,90	1,217	491,68	1,229	533,30

Pour l'exemple envisagé plus haut, $\nu = 46{,}3$ et $m_{\sigma^2} = 2{,}23$. Cet exemple montre que malgré le faible effectif de l'échantillon, l'erreur est négligeable devant celle que l'on peut obtenir pour $N = 12$ si un seul caractère est étudié. Cette augmentation de la précision est due à ce que toute la totalité de l'information sur de nombreux caractères est utilisée pour évaluer un paramètre généralisé. Une faible erreur statistique est la particularité typique de tous les paramètres généralisés envisagés dans cet article.

L'intervalle de confiance pour la valeur générale de la variance généralisée $\hat{\sigma}^2$ est trouvé d'après la formule

$$\frac{\nu}{\chi^2_\nu (1-\alpha/2)}\,\sigma^2 < \hat{\sigma}^2 < \frac{\nu}{\chi^2_\nu (\alpha/2)}\,\sigma^2$$

où $1 - \alpha$ est la probabilité de confiance donnée à l'avance; $\chi^2_\nu (\alpha/2)$ et $\chi^2_\nu (1-\alpha/2)$ sont les valeurs de table de la distribution de χ^2 avec les ν degrés de liberté et les probabilités égales respectivement à $\alpha/2$ et $1 - \alpha/2$. Soit pour notre exemple, une probabilité de confiance égale à 0,9. Alors $\alpha = 0{,}10$. A partir de la table de distribution de χ^2 nous avons pour $\nu = 46{,}3$: $\chi^2_\nu (\alpha/2) \approx 84{,}3$; $\chi^2_\nu (1 - \alpha/2) \approx 31{,}7$. Etant donné que $\sigma^2 \approx 10{,}72$, la valeur générale de la variance généralisée avec une probabilité 0,9 se trouve dans l'intervalle de confiance (5,89; 15,66).

Les variances d'échantillonnage généralisées peuvent être comparées entre elles à l'aide du critère F. Soient σ^2_1 et σ^2_2 les variances généralisées avec le nombre de degrés de liberté égal respectivement à ν_1 et ν_2. On établit le *critère de Fischer*:

$$F = \frac{\sigma^2_1}{\sigma^2_2} \quad (\text{si } \sigma^2_1 > \sigma^2_2).$$

La signification des différences est ordinairement évaluée d'après la table de distribution de F avec le nombre de degrés de liberté du numérateur et du dénominateur égal respectivement à ν_1 et ν_2. Il importe de noter que les variances généralisées comparées peuvent être déterminées d'après les complexes à nombre différent de caractères.

On peut établir la moyenne des variances généralisées. Soient σ^2_1, σ^2_2, . . ., σ^2_k des variances estimées d'après k ensembles des caractères (échantillons). Soient ν_1, ν_2, . . ., ν_k les degrés de liberté correspondants. La variance moyenne pondérée est égale à $\sigma^2 = \frac{1}{\nu}\sum \nu_i \sigma^2_i$; le nombre de degrés de liberté $\nu = \nu_1 + \nu_2 + \ldots + \nu_k$. Pour tester l'homogénéité des variances, on peut utiliser le critère de Bartlett.

Coefficient de variation généralisé

Définissons le coefficient de variation généralisé C_v comme suit :

$$C_v = \frac{\sigma}{M} \tag{4}$$

où $M = \sqrt[p]{\bar{x}_1 \bar{x}_2 \ldots \bar{x}_p}$ est la moyenne géométrique des valeurs moyennes des caractères étudiés.

On peut définir le coefficient de variation d'une autre façon :

$$C'_v = \sqrt{\frac{p}{\bar{x} S^{-1} \bar{x}^T}} \tag{4'}$$

où S^{-1} est la matrice inverse de S,

$\bar{x}$ le vecteur ligne des valeurs moyennes des caractères.

Coefficient de corrélation généralisé

L'indice caractérisant l'intensité de la liaison entre tous les caractères est introduit sur la base du déterminant de la matrice de corrélations $d = \det(R)$. On peut le définir comme $\sqrt[p]{d}$ ou comme $1 - \sqrt[p]{d}$. A notre avis, l'indice le plus commode est :

$$\rho = 1 - c \tag{5}$$

où c est la racine de l'équation

$$pc^{p-1} - (p-1)\, c^p = d, \quad 0 \leqslant c \leqslant 1 \tag{5'}$$

(v. *Annexe mathématique*). Cette racine est aisément trouvée par itérations :

$$c_i = \left(\frac{d}{p - (p-1)\, c_{i-1}} \right)^{1/p-1}, \quad c_0 = 0.$$

Pratiquement, cinq itérations suffisent. La grandeur ρ est une certaine moyenne des valeurs moyennes de toutes les corrélations paires possibles et se trouve plus près de la maximale d'entre elles. Pour des effectifs N de l'échantillon proches de p, l'estimation de la grandeur ρ peut s'écarter en tendant vers les valeurs plus grandes. Nous ne connaissons pas de correction de cet écart. Une expression pour l'erreur d'échantillonnage m_ρ n'est pas non plus connue.

Indices généralisés de l'intensité de l'influence du facteur

Soient k groupes d'individus et un groupe numéro i avec n_i individus. Soit x_{ij} le vecteur p-dimensionnel de l'individu j appartenant au groupe i ; $i = 1, 2, \ldots, k$; $j = 1, 2, \ldots, n_i$. Désignons l'effectif total par N : $N = n_1 + n_2 + \ldots + n_k$.

D'une façon analogue à l'analyse de variance unifactorielle pour le cas d'un seul caractère introduisons des matrices de covariances: la matrice intergroupe S_1 et la matrice intragroupe S_2, et la grandeur n_0:

$$S_1 = \frac{1}{k-1}\left(\sum_{i=1}^{k} n_i \bar{x}_i^T \bar{x}_i - N\bar{x}^T\bar{x}\right),$$

$$S_2 = \frac{1}{N-k}\left(\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}^T x_{ij} - \sum_{i=1}^{k} n_i \bar{x}_i^T \bar{x}\right), \qquad (6)$$

$$n_0 = \frac{1}{k-1}\cdot\left(N - \frac{1}{N}\sum_{i=1}^{k} n_i^2\right)$$

où $\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$ et $\bar{x} = \frac{1}{N}\sum_i\sum_j x_{ij}$ sont les vecteurs des valeurs moyennes du groupe i et de la moyenne générale.

Désignons par $\hat{S}_w$ et $\hat{S}_b$ les valeurs générales des matrices de covariances intragroupe (« within ») et integroupe (« between »). De façon analogue au cas où $p = 1$, nous avons (M est le symbole de l'espérance mathématique):

$$\mathrm{M}\{S_1\} = n_0\hat{S}_b + \hat{S}_w, \quad \mathrm{M}\{S_2\} = S_w.$$

Par conséquent, l'estimation d'échantillonnage des matrices de covariances aléatoire et factorielle est:

$$S_b = \frac{1}{n_0}(S_1 - S_2), \quad S_w = S_2. \qquad (7)$$

Pour analyser la structure de ces matrices, on peut utiliser la méthode des composantes principales ou les méthodes de l'analyse factorielle. Nous n'envisagerons pas ici les questions qui peuvent alors surgir, nous passerons à la discussion de savoir quels paramètres généralisés caractérisent la variabilité dans le groupe et entre les groupes.

Supposons pour l'instant que les valeurs générales des matrices de covariances $\hat{S}_w$ et $\hat{S}_b$ soient connues, posons $\hat{S} = \hat{S}_b + \hat{S}_w$. Trouvons les valeurs générales des variances généralisées: $\hat{\sigma}^2 = \sqrt[p]{\det(\hat{S})}$, $\hat{\sigma}_w^2 = \sqrt[p]{\det(\hat{S}_w)}$, $\hat{\sigma}_b^2 = \sqrt[p]{\det(\hat{S}_b)}$. Dans le cas où $p = 1$, on a toujours $\hat{\sigma}^2 = \hat{\sigma}_b^2 + \hat{\sigma}_w^2$. Cependant, si $p > 1$, on a toujours $\hat{\sigma}^2 > \hat{\sigma}_b^2 + \hat{\sigma}_w^2$. Il n'y aura égalité que si les matrices $\hat{S}_b$ et

$\hat{S}_w$ sont semblables. La similitude biologique de ces matrices signifie que le facteur en question agit de façon isotrope sur les caractères étudiés, c'est-à-dire rigoureusement dans le sens de la variabilité dans les groupes. Si l'action du facteur n'est pas isotrope, c'est-à-dire que si elle ne coïncide pas tout à fait avec la direction de la variabilité dans les groupes, il n'y aura pas d'égalité. Dans ce cas, en plus de $\hat{\sigma}_b^2$ et $\hat{\sigma}_w^2$, on introduit encore une composante de la variance généralisée (composante d'anisotropie) définie comme

$$\hat{\sigma}^2 - \hat{\sigma}_b^2 - \hat{\sigma}_w^2.$$

Ainsi, la structure de l'influence du facteur sur l'ensemble des caractères peut être décrite par deux paramètres généralisés :

1° *l'indice généralisé d'intensité de l'influence du facteur ou le coefficient généralisé de corrélation entre classes* (*indice de différence entre groupes*)

$$\hat{r}_w = 1 - \frac{\hat{\sigma}_w^2}{\hat{\sigma}^2}; \tag{8}$$

2° *l'indice généralisé d'isotropie de l'action du facteur*

$$\hat{r}_I = \frac{\hat{\sigma}_b^2}{\hat{\sigma}^2 - \hat{\sigma}_w^2}. \tag{9}$$

Ce dernier indice est un paramètre essentiellement nouveau ressortant lors de l'étude de l'ensemble des caractères comme un tout. Il indique le caractère de la divergence des groupes par rapport à la variabilité dans les groupes. Si $\hat{r}_I < 1$, ceci signifie que la direction de la divergence des groupes ne coïncide pas avec celle de la variabilité à l'intérieur des groupes.

Si les groupes étudiés sont des groupes de parents à un certain degré, la grandeur r_w est interprétée génétiquement de la même façon que dans le cas d'un seul caractère. Si $r_I < 1$, la direction de la variabilité génotypique ne coïncide pas avec celle de la variabilité phénotypique.

Passons maintenant aux questions de l'évaluation pratique des indices mentionnés. Introduisons la matrice de covariances S définie par toutes les données de départ N. Pour simplifier les calculs, nous recommandons la formule

$$S = \frac{1}{N-1}\left[(k-1)\,S_1 + (N-k)\,S_2\right]$$

si les matrices S_1 et S_2 sont déjà connues grâce aux formules (6). Les valeurs de $\hat{\sigma}^2_w$ et de $\hat{\sigma}^2$ sont évaluées d'après les formules

$$\sigma^2_w = G \sqrt[p]{\det (S_2)} \text{ et } \sigma^2 = G \sqrt[p]{\det (S)} \tag{10}$$

où les corrections d'écart $G_{p,N}$ ne sont pas les mêmes pour les deux variances et sont cherchées dans le *Tableau 1* avec les paramètres p, $N - k + 1$ et p et N.

En partant de l'*Annexe mathématique*, l'estimation de l'influence généralisée du facteur est donnée par la formule

$$r_w = \frac{1-y}{1-\theta y} \tag{11}$$

où $y = \frac{\sigma^2_w}{\sigma^2}$, $\theta = 1 - \frac{n_0 (k-1)}{N-1}$.

La variance σ^2_b pourrait être évaluée comme $\sqrt[p]{\det (S_b)}$. Cependant des difficultés surgissent, elles sont dues à l'absence de précision positive de l'estimation d'échantillonnage de la matrice S_b. Nous proposons en l'occurrence l'évaluation suivante (v. *Annexe mathématique*) :

$$\sigma^2_b = \sigma^2_1 \left(1 - \frac{z}{p}\right) \tag{12}$$

où $\sigma^2_1 = \frac{G}{n_0} \sqrt[p]{\det (S_1)}$,

G est cherchée dans la table avec les paramètres p et k,
z est la somme des éléments diagonaux de la matrice $S_1^- S_w$.
L'indice d'isotropie est trouvé d'après la formule

$$r_I = \frac{1-r_w}{r_w} \cdot \frac{\sigma^2_b}{\sigma^2_w}. \tag{13}$$

Exemple. N. Glotov et L. Sémérikov [2, 8] ont étudié deux espèces de chêne — le chêne pédonculé (*Quercus robur* L.) et le chêne oriental (*Q. macranthera* Fisch. et Mey) — dans quatre localités de Daghestan. Chacune des populations étudiées était représentée par plusieurs arbres (10 et plus). On prélevait sur chaque arbre 5 feuilles intactes (une talle). Les caractères suivants de la feuille étaient analysés : longueur de la feuille, longueur du pétiole, hauteur de la pointe, largeur des lobes, largeur des découpures, et certains autres. Le *Tableau 2* présente les résultats de l'analyse généralisée de la variabilité des cinq premiers caractères dans quatre populations.

Nous n'envisagerons pas ici les données obtenues. Notons seulement que les paramètres généralisés révèlent nettement les particularités des espèces comparées.

Tableau 2

Paramètres généralisés de la variabilité selon cinq caractères définissant les dimensions de la feuille pour deux espèces de chêne

Espèce	*Q. macranthera*		*Q. robur*	
Localité	Vanachi	Bourtounaï	Guerga	Pokrov
Variance à l'intérieur de l'arbre (σ_w^2)	41,0	32,6	12,5	8,5
Coefficient de variation (C_v')	37,6	36,1	15,3	15,1
Coefficient de corrélation (ρ)	0,68	0,72	0,49	0,51
Indice de différence entre les arbres (r_w)	0,34	0,27	0,73	0,70
Indice d'isotropie (r_I)	~ 0,42	~ 0,31	0,96	0,98

Indice de similitude des populations

Désignons par $\bar{x}_1$ le vecteur des valeurs moyennes et par S_1 la matrice de covariances trouvés pour un échantillon d'effectif N_1 tiré de la première population. Soient $\bar{x}_2$ le vecteur des valeurs moyennes et S_2 la matrice de covariances obtenus pour un échantillon d'effectif N_2 tiré de la deuxième population. Posons $S = {}^1/_2 (S_1 + S_2)$.

Comme mesure de la proximité des populations on utilise généralement la distance de Mahalanobis [4] désignée par D^2 et calculée d'après la formule

$$D^2 = (\bar{x}_1 - \bar{x}_2)\, S^{-1}\, (\bar{x}_1 - \bar{x}_2)^T. \qquad (14)$$

Notons un défaut de cette mesure de la proximité. En partant de cette mesure, la distance entre populations est égale à zéro dans le cas où les populations ne diffèrent pas par les moyennes des caractères. Cependant, ce n'est pas toujours logique. Par exemple, les populations comparées peuvent avoir des moyennes identiques pour tous les caractères, mais se distinguer l'une de l'autre par les variances de ces caractères ou par les valeurs des corrélations entre elles. Les deux cas peuvent être dus à des différences biogéocénotiques de l'environnement de ces populations, à leur structure génétique diverse et, pour cette raison, de telles populations ne sauraient être considérées comme semblables. On ne peut donc utiliser la mesure D^2 pour les comparer, car en l'occurrence elle n'est pas adéquate au problème.

Il existe des mesures tenant compte d'une inégalité possible des matrices de covariances. Telle est celle de S. Kullback [6] reposant

sur des principes de l'information statistique. Dans cet article, une autre approche est développée dont les principes sont les mêmes pour les caractères aussi bien quantitatifs que qualitatifs. Ceci permet de comparer de façon adéquate la similitude entre les populations pour différents caractères. La méthodologie de l'approche a déjà été exposée par l'auteur [3].

Soit r l'indice de similitude entre deux populations pour un ensemble de p caractères. Désignons par

$$r_x = e^{-\frac{D^2}{8p}} \tag{15}$$

la part de la similitude résultant de différences entre les moyennes. Si les moyennes coïncident, $r_x = 1$, sinon $r_x < 1$. Soit σ_i^2 la variance généralisée pour la population i, $i = 1, 2$. La valeur de $G_{p.N_i}$ est alors cherchée dans le *Tableau 1* avec les paramètres p et N_i. Soit $\sigma^2 = G_{p,N} \sqrt[p]{\det (S)}$ où les valeurs des paramètres pour G sont égales à p et à $N = N_1 + N_2$.

Désignons par

$$r_S = \frac{\sqrt{\sigma_1 \sigma_2}}{\sigma} \tag{16}$$

la part de la similitude résultant de différences entre les matrices de covariances des populations comparées. Si les matrices coïncident, $r_S = 1$, sinon $r_S < 1$.

L'indice de similitude tenant compte aussi bien de différences entre les moyennes que de différences entre les matrices de covariances est défini comme

$$r = r_x r_S. \tag{17}$$

Notons que cet indice tient compte de l'exigence mentionnée plus haut : il est normalisé par rapport au nombre de caractères. La déduction de la formule est présentée dans l'*Annexe mathématique*.

Il est intéressant de remarquer que cette approche fondée sur les conditions de définition de l'indice de similitude [3] et sans connexion par l'idée avec les raisonnements conduisant à la mesure D^2 est tout de même reliée à la distance de Mahalanobis. Si les matrices de covariances étaient semblables, la grandeur r coïnciderait avec D^2 à une transformation logarithmique près. Cependant, l'indice r tient compte aussi de la différence de ces matrices, il est donc plus exact. Comme pour les caractères polymorphes, l'indice r évalue la part de la variation totale des caractères dans les populations comparées, le degré de transgression de distributions des caractères.

Revenons à l'exemple examiné plus haut. Le *Tableau 3* représente les données sur la similitude entre quatre populations de *Q. robur* au Daghestan d'après cinq caractères de la feuille : longueur de la feuille,

Tableau 3

Indices de similitude r (moitié supérieure du tableau) et r_x (moitié inférieure) pour l'ensemble de 5 caractères de 4 populations du chêne (*Q. robur*) au Daghestan

Population	Pokrov	Dylyme	Bouïnaxk	Guerga
Pokrov		0,95	0,85	0,91
Dylyme	0,98		0,89	0,86
Bouïnaxk	0,92	0,94		0,78
Guerga	0,94	0,92	0,86	

longueur du pétiole, hauteur de la pointe, largeur des lobes, largeur des découpures. La moitié inférieure du tableau reflète la similitude, déterminée seulement par la mesure de Mahalanobis (r_x) ; la moitié supérieure, l'indice de similitude r.

Le *Tableau 3* montre que l'indice r peut différer significativement de la mesure tenant compte seulement de la différence entre les r_x moyens et que son pouvoir différenciant est plus élevé. De plus, le membre r_S comprend les différences entre les populations comparées non seulement d'après les variances, mais aussi d'après les coefficients de corrélation. Par exemple, les populations Dylyme et Bouïnaxk ont des variances généralisées presque identiques : pour Dylyme $\sigma^2 = 21{,}9$, pour Bouïnaxk $\sigma^2 = 22{,}0$. Cependant, leurs matrices de corrélations diffèrent fortement (*Tableau 4*), ce qui a permis une meilleure distinction des populations en utilisant l'indice r.

Tableau 4

Matrice de corrélations pour les populations Dylyme (moitié supérieure du tableau) et Bouïnaxk (moitié inférieure)

Caractères	Longueur de la feuille	Longueur du pétiole	Hauteur de la pointe	Largeur des lobes	Largeur des découpures
Longueur de la feuille	1	0,56	0,73	0,84	0,56
Longueur du pétiole	—0,15	1	0,43	0,61	0,63
Hauteur de la pointe	0,61	—0,11	1	0,63	0,28
Largeur des lobes	0,49	—0,05	0,40	1	0,64
Largeur des découpures	0,30	0,02	0,28	0,29	1

En ce qui concerne l'erreur statistique de l'indice de similitude d'après les caractères quantitatifs on n'est pas arrivé à lui trouver une formule simple, à la différence du cas des caractères polymorphes. Des considérations heuristiques conduisent à cette formule approchée pour l'erreur statistique m_r:

$$m_r = \frac{1}{2pr^{p-1}} \sqrt{\frac{N_1+N_2}{N_1N_2}(1-r^{2p})}, \tag{18}$$

où N_1et N_2 sont les effectifs des échantillons des populations comparées. Il ne faut pas oublier que cette formule donne une valeur trop élevée de l'erreur.

Admettons, par exemple, que nous ayons obtenu pour $p = 5$, $r = 0{,}92$. Soit $N_1 = 30$ et $N_2 = 50$, alors $m_r \approx 0{,}024$.

Indice généralisé de similitude de plusieurs populations

Supposons qu'il y ait k échantillons. Leurs effectifs sont N_1, N_2, . . ., N_k. Admettons que pour chaque échantillon soient évalués les vecteurs des moyennes $\bar{x}_1$, $\bar{x}_2$, . . ., $\bar{x}_k$ et que leurs matrices de covariances soient S_1, S_2, . . ., S_k. Dans la suite, nous aurons besoin de matrices A_j inverses des matrices de covariances: $A_j = S_j^{-1}$. Introduisons les grandeurs suivantes:

$$A = \frac{1}{k}\sum_{j=1}^{k} A_j; \quad \bar{x} = \frac{1}{k} A^{-1}\left(\sum_{j=1}^{k} A_j \bar{x}_j^T\right);$$

$$H^2 = \frac{4}{k}\sum_{j=1}^{k} (\bar{x}_j - \bar{x})\bar{A}_j(\bar{x}_j - \bar{x})^T.$$

Soit

$$R_S = \frac{\sqrt[k]{\sigma_1\sigma_2 \ldots \sigma_k}}{\sigma}, \quad R_x = e^{-\frac{H^2}{8p}}. \tag{19}$$

Alors, l'indice de similitude de plusieurs populations est trouvé d'après la formule

$$R = R_x R_S. \tag{20}$$

L'indice R est un r généralisé pour le cas d plusieurs échantillons et est introduit par analogie avec le cas des caractères polymorphes [3].

A propos, la grandeur H^2 est la distance de Mahalanobis généralisée pour le cas où le nombre d'échantillons k est supérieur à 2.

DISCUSSION

Les paramètres généralisés introduits plus haut sont des indices intégraux de la variabilité des populations selon l'ensemble des caractères quantitatifs, « imprégnés » de toutes les particularités de la variabilité de chaque caractère pris à part. L'approche d'une étude de la variabilité des populations d'après un complexe de caractères, en utilisant un nombre réduit de paramètres intégraux, ne rejette pas les approches rattachées à l'étude des caractères particuliers. Au contraire, il est préférable d'analyser d'abord les caractères isolés ou leurs indices et, ensuite, de les diviser en un certain nombre de groupes, par exemple au moyen de l'échantillonnage par clusters ou d'une autre méthode. Si le groupe de caractères isolé est unifié en un certain sens (s'il représente, par exemple, une unité structurale déterminée), il est naturel de chercher pour ce groupe un paramètre caractérisant une unité structurale donnée. Par exemple, les cinq caractères de la feuille mentionnés plus haut définissent ses dimensions et en partie sa forme, et, ici, les paramètres généralisés caractérisent la variabilité de la feuille en tant que feuille et non pas celle de ses caractères particuliers. Il faut noter l'avantage important des paramètres généralisés. Ils ont une grande précision de l'estimation (une faible erreur statistique) même pour des échantillons relativement petits du fait que toute la quantité d'information sur un grand nombre de caractères se concentre dans un seul paramètre généralisé. Les formules pour calculer les erreurs statistiques sont assez simples. A notre avis, l'approche d'une étude de la variabilité des populations basée sur des paramètres généralisés peut, outre la méthode des composantes principales et l'analyse factorielle, aider à étudier la structure des populations d'après un ensemble de caractères quantitatifs.

ANNEXE MATHÉMATIQUE

1° **Déduction des formules pour G et ν.** Soient S et $\hat{S}$ des matrices de covariances d'échantillonnage et générale. Soit $s^2 = \sqrt[p]{\det(S)}$, $\hat{\sigma}^2 = \sqrt[p]{\det(\hat{S})}$. Posons $\xi = \frac{s^2}{\hat{\sigma}^2}(N-1)$, $M\,\xi^t = m_t$. On sait [1] que

$$m_t = 2^t \prod_{j=1}^{p} \frac{\Gamma\left(\frac{t}{p} + \frac{N-j}{2}\right)}{\Gamma\left(\frac{N-j}{2}\right)}.$$

On sait également que m_p est distribuée comme le produit de p grandeurs indépendantes aléatoires qui suivent une distribution χ^2. Trouvons un nombre g tel que, pour p et N donnés, la grandeur $g\xi$

soit distribuée à peu près comme χ^2 avec ν degrés de liberté. Posons égaux les deux premiers moments de distribution: $gm_1 = \nu$, $g^2m_2 = \nu(\nu + 2)$. Nous continuerons à chercher une $G_{p,N}$ telle que $\mathbf{M}(Gs^2) = \sigma^2$. Nous obtiendrons alors:

$$G = \frac{N-1}{m_1}, \quad \nu = 2\frac{m_1^2}{m_2 - m_1^2}. \tag{21}$$

Ces formules ont été utilisées pour dresser le *Tableau 1*. En partant de la formule de Stirling, nous arrivons à l'approximation suivante:

$$G = \frac{N}{N-1}\exp\left\{\frac{p}{2N} + \frac{p^2}{6N^2}\right\}, \quad \nu = \frac{p^2}{\ln\left(\frac{N-1}{N-p-1}\right)}. \tag{22}$$

Elle est beaucoup plus précise que l'estimation donnée par Hoel [1].

2° **Déduction de la formule pour la corrélation généralisée ρ.** La formule est déduite à partir des considérations que si tous les coefficients de corrélation paire sont égaux à une même grandeur r, le déterminant de la matrice de corrélations sera $\det(R) = pc^{p-1} - (p-1)c^p$, $c = 1-r$. Par conséquent, l'indice ρ est une certaine moyenne de tous les coefficients de corrélation proche du plus grand d'entre eux.

3° **Déduction de la formule pour r_w et r_I.** Nous utilisons l'équation qui se déduit de la formule de la dérivée du logarithme du déterminant [6]:

$$\frac{d^k}{d\theta^k}(\ln(\det(A+\theta B))^{1/p})_{\theta=0} = (-1)^{k-1}\frac{(k-1)!}{p}\operatorname{tr}(A^{-1}B)^k$$

Admettons que les matrices A et B soient semblables ($A \sim B$), c'est-à-dire qu'il existe un nombre r tel que $B = rA$. Il en résulte: $r = \frac{\sqrt[p]{\det(B)}}{\sqrt[p]{\det(A)}}$. De la formule déduite plus haut pour la dérivée il ressort que $1/p \ln(\det(A + \theta B)) = 1/p \ln(\det(A)) + \ln(1 + r\theta) + o(\theta)$, d'où, pour de petites valeurs du paramètre θ et des matrices A et B semblables, on obtient une égalité approchée

$$\sqrt[p]{\det(A+\theta B)} \approx (1+r\theta)\sqrt[p]{\det(A)}.$$

Selon la méthode standard [9], on peut montrer que $\mathbf{M}\{S_1\} = n_0\hat{S}_b + \hat{S}_w$, $\mathbf{M}\{S_2\} = \hat{S}_w$ où M est le signe de l'espérance mathématique, S_1, S_2 et n_0 sont trouvés d'après la formule (6). Si S est l'estimation d'échantillonnage de la matrice de covariances générale, déterminée pour toutes les observations N, $\mathbf{M}\{S\} = \hat{S}_w + \alpha\hat{S}_b$ où $\alpha = \frac{n_0(k-1)}{N}$.

Posons $A = \hat{S}_w + \alpha\hat{S}_b$, $B = \hat{S}_w$, $\theta = 1 - \alpha$. Alors $A - \theta B = \alpha\hat{S}$ où $\hat{S} = \hat{S}_w + \hat{S}_b$. Pratiquement, la part B dans A est suffisamment grande et donc $B \sim A$. Aussi peut-on utiliser la formule déduite plus haut. Nous avons:

$$\alpha \sqrt[p]{\det(\hat{S})} = (1-\theta y) \sqrt[p]{\det(A)}, \quad r_e = \frac{\alpha\hat{y}}{1-\theta\hat{y}}$$

où $\hat{y} = \dfrac{\sqrt[p]{\det(\hat{S}_w)}}{\sqrt[p]{\det(A)}}$, $\hat{r}_e = \dfrac{\sqrt[p]{\det(\hat{S}_w)}}{\sqrt[p]{\det(\hat{S})}}$.

Ayant posé $\hat{r}_w = 1 - \hat{r}_e$, nous trouvons $\hat{r}_w = \dfrac{1-\hat{y}}{1-\theta\hat{y}}$. Remplaçons maintenant les valeurs générales des paramètres par les valeurs échantillonnées. Nous obtenons la formule (11). La formule (12) se déduit de la formule de la dérivée du logarithme du déterminant. La signification de r_w peut être évaluée en se basant sur y [4]. Le critère statistique pour évaluer r_I n'est pas connu.

4° **Indices de similitude.** Par analogie avec l'indice de similitude d'après les caractères polymorphes [3], nous introduisons des paramètres semblables également pour les distributions continues. Soient $q_1(x)$ et $q_2(x)$ les densités de distribution d'un vecteur p-dimensionnel x de deux populations comparées. Déterminons la similitude entre elles comme

$$r = \left(\int_{-\infty}^{\infty} \sqrt{q_1(x)\, q_2(x)}\, dx \right)^{1/p}.$$

Si $q_1, q_2, \ldots, q_k$ sont les densités de k populations comparées, la similitude générale de toutes les k populations est définie comme

$$R = \left(\int_{-\infty}^{\infty} \sqrt[k]{q_1(x) \ldots q_k(x)}\, dx \right)^{1/p}$$

en accord avec les conditions imposées auparavant.

Bhattacharya (1943, cité d'après Kullback [6]) avait proposé l'indice r comme mesure de la similitude d'une paire de populations, mais par la suite il a utilisé la transformation de cette grandeur pour le cas des caractères polymorphes.

Si les densités $q_i(x)$ sont normales:

$$q_i(x) = \frac{1}{\sqrt{\det(2\pi\hat{S}_i)}} \exp\left\{ -\frac{1}{2}(x-\hat{x}_i)\, \hat{S}_i^{-1}\, (x-\hat{x}_i)^T \right\},$$

les expressions pour r et R sont déterminées d'après les matrices de covariances et les vecteurs des moyennes. En intégrant ces expressions, on parvient aux formules (15)-(17) et (19)-(20).

RÉFÉRENCES

[1] *Anderson T. W.* An introduction to multivariate statistical analysis. London, Chapman and Hall, 1958.
[2] *Глотов Н. В., Семериков Л. Ф.* Изменчивость дуба черешчатого (*Quercus robur* L.) в Дагестане. В сб.: Проблемы эволюционной и популяционной генетики, Махачкала, Даг. фил. АН СССР, 1978, 78-85.
[3] *Животовский Л. А.* Показатель сходства популяций по полиморфным признакам. *Журн. общ. Биол.*, 1979, 40, 4, 587-602.
[4] *Kendall M. G., Stuart A.* Design and analysis, and time-series. London, Charles Griffin, 1974.
[5] *Cramer H.* Mathematical methods of statistics. Stokholm, 1946.
[6] *Kullback S.* Information theory and statistics. N.Y., J. Wiley, 1959.
[7] *Lawley D., Maxwell A.* Factor analysis as a statistical method. London, 1963.
[8] *Семериков Л. Ф., Глотов Н. В.* Изменчивость дуба восточного (*Quercus macranthera*, Fisch. et Mey) в Дагестане. В сб.: Популяции растений, Л., ЛГУ, 1979, 179-189.
[9] *Snedecor G. W., Cochran W. G.* Statistical methods. Ames, 1967, 593 p.

L'ENVIRONNEMENT ET LE FARDEAU GÉNÉTIQUE DES POPULATIONS HUMAINES

PROBLÈMES GÉNÉTIQUES DE LA POLLUTION DE L'ENVIRONNEMENT

R. Khiltchevskaïa

Les *polluants mutagènes* de la biosphère sont divisés en trois classes: physiques (radiation ionisante dont l'effet mutagène a été le plus complètement étudié), biologiques (les virus par exemple) et chimiques qui représentent à cause de leur abondance le principal danger.

Tous les *mutagènes chimiques* sont classés en trois groupes principaux [21]: les substances inorganiques naturelles (oxydes d'azote, nitrates, nitrites, plomb, cadmium, corps radioactifs, etc.) et les substances naturelles d'origine organique (alcaloïdes, hormones, etc.); les composés naturels transformés (produits tirés du pétrole, produits de la combustion du charbon et du bois, composés des métaux lourds, déchets alimentaires, composés chimiques largement utilisés, tels que l'uréthane, le tétrachlorure de carbone, l'oxyde d'éthylène, le chloroprène, etc.); les produits chimiques non rencontrés autrefois dans la nature (pesticides, additifs alimentaires, déchets industriels, médicaments).

Nous sommes entrés dans une époque de mutagénèse induite agissant pratiquement partout sur l'information génétique de presque toutes les espèces, l'homme compris. Ainsi, par exemple aux U.S.A., plus de 2 millions de composés chimiques (d'après des calculs approximatifs) étrangers à la biosphère avaient été déversés dans l'environnement déjà en 1971. De nos jours, l'humanité utilise jusqu'à 50 000 rien que de médicaments divers. La tendance à réduire le nombre d'années nécessaire au passage des recherches en laboratoire à la production de masse des polymères [16] montre bien les cadences toujours accélérées du développement de la production et de l'entrée dans la circulation de nouveaux composés chimiques synthétiques.

© Издательство Химия, 1979.

L'utilisation universelle des pesticides présente le plus de danger pour la santé humaine, ce que confirment les calculs des indices de stress effectués en R.F.A. [41]. L'assortiment mondial de pesticides compte actuellement plus de 100 000 dénominations [49]. Pour leur synthèse, on utilise environ 900 composés organiques et minéraux dont la plupart possèdent un effet mutagène potentiel. Un des derniers aperçus [44] sur l'étude des pesticides en tant qu'agents mutagènes du milieu ambiant montre que sur 239 composés examinés 119 exerçaient un effet mutagène, soit 49,8 %, c'est-à-dire presque la moitié des produits étudiés. On ne saurait considérer ce nombre comme définitif étant donné que l'activité génétique de la plupart d'entre eux (69,5 %) n'a été vérifiée que sur un seul objet biologique. Les données sur l'activité mutagène des pesticides procurées par plus de 100 travaux effectués dernièrement sont présentées au *Tableau 1*.

Tableau 1

Effet mutagène des pesticides [40]

Pesticides	Rapport du nombre des pesticides étudiés au nombre des pesticides produisant un effet mutagène				
	sur les microorganismes	sur les plantes	sur les insectes	sur les mammifères	Total
Fongicides	17/8	11/10	6/3	10/6	44/25 (57%)
Insecticides	16/5	23/20	8/5	23/14	70/44 (63%)
Herbicides	43/4	25/24	12/9	17/4	97/41 (42%)
Autres	1/1	10/10		8/4	19/15 (88%)
Total (taux de mutagènes découverts sur un objet-test donné)	77/18 (23%)	69/64 (93%)	26/17 (65%)	58/28 (48%)	230/125 (54%)

Pris dans l'ensemble, les facteurs mutagènes frappent les structures héréditaires de tout ce qui vit. En cas d'atteinte du matériel héréditaire des cellules somatiques, les mutagènes peuvent provoquer le cancer, réduire la durée de vie, affaiblir la protection immunitaire des organismes, abaisser la résistance générale aux maladies. Par conséquent, les résultats de leurs actions peuvent se manifester au cours de l'existence d'une génération et être dénombrés par la statistique médicale. Un autre danger plus menaçant est l'atteinte des cellules embryonnaires. Ce danger est invisible, car il n'influe pas

sur la santé de la génération présente, mais il peut se manifester dans les générations suivantes sous forme de fréquence accrue des maladies héréditaires, de malformations, de mortinatalité, de diminution de la viabilité, ce qui touche également les animaux, les plantes et l'homme.

Dans la grande majorité des cas, les mutations ont une influence néfaste sur les possibilités d'adaptation de n'importe quelle espèce biologique [19]. L'accélération de la cadence des mutations entraîne la croissance du nombre des victimes génétiques dans les populations et, par conséquent, des phénomènes de crise dans la génétique des populations. Pour surmonter ces phénomènes, des processus d'adaptation héréditaire aux mutagènes sont mis en branle, ils modifient la marche de l'évolution historiquement constituée. Un exemple en est donné par l'adaptation des espèces naturelles observée dans des expériences sur un fond élevé de radiation ionisante artificiellement créé [22].

Sous l'action des pesticides, il se produit une évolution des propriétés phénotypiques basée sur la sélection des mutants stables. L'adaptation génétique acquiert les particularités de phénomènes évolutifs impulsifs [21]. C'est ainsi qu'en 1960 déjà, on avait déterminé 138 espèces d'insectes et de tiques qui offraient une résistance acquise aux insecticides chlorés et organophosphorés [36]. Actuellement, on a découvert une résistance aux poisons chimiques chez 250 espèces de parasites agricoles et de vecteurs de maladies humaines.

Les différences dans les systèmes biologique et génétique des populations déterminent la réussite de l'adaptation génétique des espèces [21]. Les espèces ayant un grand effectif et une succession rapide des générations sont les plus favorisées. La cadence accélérée de la reproduction et la courte durée d'existence de telles populations assurent la rapidité des processus d'adaptation génétique et contribuent à faire apparaître de nouvelles formes, notamment dans le monde des bactéries et des virus (pathogènes du nombre), résistantes aux produits médicamenteux. Au contraire, les petites populations à durée de vie prolongée subissent de lourdes pertes jusqu'à l'extinction complète de certaines espèces. Tous ces processus entraînent un remaniement de la structure des biocénoses et une reconstruction des écosystèmes dans le sens de leur simplification [14].

Cependant, si l'adaptation génétique de toutes les formes organiques s'accompagne d'un nombre énorme de morts de victimes génétiques, balayées par la sélection, pour l'Homme, la sélection a perdu son sens passé de facteur dirigeant l'évolution. Maintenant, la sélection ne retient la propagation dans les populations que des mutations qui sont incompatibles avec la vie et éliminées aux stades précoces de l'embryogénèse ou qui sont responsables de la naissance de porteurs de lourdes tares héréditaires. Toutes les autres modifications

de l'information génétique vont se verser dans la gigantesque variabilité génétique non sélective avec un fardeau continûment croissant de mutations récessives (latentes) [26] qui, la pollution de l'environnement par des mutagènes allant en augmentant, peut entraîner, à une étape déterminée, un dommage irréparable au programme génétique de l'espèce *Homo sapiens.*

Ainsi, l'adaptation génétique des populations humaines à une pollution toujours accrue de l'environnement par des mutagènes s'avère impossible et des mesures immédiates sont nécessaires aussi bien pour empêcher l'afflux de mutagènes biologiquement actifs dans le milieu ambiant que pour déterminer la vitesse de la mutation dans les populations humaines. La prévision des séquelles génétiques de la pollution de l'environnement par les produits de l'activité économique de l'homme est indispensable.

La solution des problèmes qui se posent à la génétique suit deux directions principales : la création de systèmes testants sensibles pour mettre en évidence et évaluer les mutagènes de l'environnement — substances polluant la biosphère, et la recherche d'approches de la réalisation d'un monitoring de la dynamique du fardeau génétique dans les populations humaines. La troisième direction sera l'élaboration d'un complexe de mesures pour l'utilisation rationnelle des ressources naturelles de la biosphère, basée sur l'étude de la structure génétique des populations naturelles et agricoles.

Élaboration de systèmes testants sensibles pour détecter les mutagènes

Les premières recherches sur la mutagénèse induite (provoquée) peuvent être rapportées aux années 1920 (début de l'étude de l'effet mutagène de la radiation ionisante et des rayons UV). Dans les années 1940-1950, on a entrepris les recherches sur la mutagénèse chimique et, en 1970, d'outil expérimental les mutagènes de nature physique et chimique sont devenus un élément constant de la biosphère qui, associé à d'autres facteurs, agit sur l'hérédité des organismes vivants [20].

Le dénombrement des troubles génétiques provoqués par des mutagènes de différente nature se fait à trois niveaux différents : mutations au niveau du génome (dans le noyau cellulaire le nombre des chromosomes varie), mutations au niveau des chromosomes (la structure du chromosome se modifie), mutations au niveau du gène (des perturbations se produisent dans la molécule d'ADN du gène). Les mutations récessives ayant des effets nocifs faiblement prononcés représentent la classe de mutations la plus répandue. Elles sont dues à l'action prolongée des mutagènes chimiques de l'environnement et présentent probablement le plus grand danger potentiel. Suivant le mécanisme de leur action, les mutagènes chimiques se

divisent en neuf classes principales [20]: composés alcoylants; peroxydes; aldéhydes; hydroxylamines; composés de l'azote; antimétabolites et, du nombre, les analogues des bases d'ADN; sels des métaux lourds; colorants basiques; substances aromatiques (certains cancérogènes et alcaloïdes, certains médicaments, les pesticides).

Afin de créer des systèmes testant l'effet mutagène des polluants chimiques, on a naturellement utilisé en premier lieu les tests destinés à l'étude de l'activité mutagène d'une radiation ionisante. La grande ressemblance entre les résultats de l'action d'une radiation ionisante et des composés chimiques a même poussé à essayer d'utiliser des équivalents de dose de la radiation ionisante pour les mutagènes chimiques [11, 17]. Mais il s'est avéré que la mutagénèse chimique avait ses particularités spécifiques [20, 28, 29].

Le fonctionnement parfait des systèmes testant l'effet mutagène suppose qu'ils répondent à certaines conditions [20].

1° Le système testant doit informer: sur la fréquence et les types des mutations ponctuelles (intragéniques) liées à la transition et à la transversion des paires de bases et à des délétions à l'intérieur du gène; sur la fréquence et les types de remaniements structuraux des chromosomes; sur les modifications dans le nombre des chromosomes. Pour l'instant, un seul objet, la drosophile, satisfait à toutes les conditions imposées.

2° La sensibilité et la reproductivité du système testant sont ses éléments essentiels sans lesquels il est impossible de tester les mutagènes en faibles doses. Les systèmes testants bactériens tiennent la première place pour ces indices.

3° Pour apprécier les effets des mutagènes sur les mammifères, il faut tenir compte des modifications métaboliques (il existe plusieurs travaux indiquant que certaines substances inertes au point de vue mutagène deviennent actives et *vice versa*, en pénétrant dans la voie du métabolisme des mammifères). Il est possible de remplir cette exigence en utilisant des systèmes testants bactériens avec activation métabolique tant *in vitro* qu'*in vivo*.

4° Les frais causés par les épreuves ne doivent pas être trop élevés.

5° La méthode doit autant que possible exiger peu de temps (pour l'instant, seuls les tests utilisant des microorganismes le permettent).

Actuellement, on a proposé plus d'une trentaine de systèmes testants faisant face sinon à toutes les conditions formulées, du moins à une partie d'entre elles. Chaque année, le nombre de systèmes testants va en augmentant, on cherche à perfectionner ceux dont on dispose déjà (*Tableau 2*).

Les systèmes testants les plus répandus dans la pratique de la détermination des propriétés mutagènes des substances chimiques sont décrits en détail dans plusieurs travaux [1, 8, 12, 29, 30, 31, 34, 37, 39, 45, 46, 54, 58, 59, 62].

Tableau 2

Appréciation de la valeur et de l'efficacité des systèmes testants les plus souvent utilisés dans le screening des mutagènes

Système testant	Durée de l'épreuve	Valeur *		Facilité relative de la détection	
		Frais courants	Frais de mise en exploitation	Mutations géniques	Aberrations chromosomiques
Microorganismes à activation métabolique					
Salmonella typhi murium	2 à 3 jours	Très bas	Bas	Excellente	
Escherichia coli	2 à 3 jours	*Idem*	*Idem*	*Idem*	
Levure	3 à 5 jours	*Idem*	*Idem*	Bonne	Inconnue
Neurospora crassa	1 à 3 semaines	Moyens	Moyens	Très bonne	Bonne
Cellules cultivées de mammifères à activation métabolique	2 à 5 semaines	De moyens à élevés **	*Idem*	Excellente	Inconnue
Milieu du receveur					
avec microorganismes	2 à 7 jours	De bas à moyens	De bas à moyens	Bonne	
avec cellules de mammifères	2 à 5 semaines	De moyens à élevés	Moyens	Inconnue	Bonne
Plantes					
Vicia faba	3 à 8 jours	Bas	Bas		Peu claire
Tradescantia paludosa	2 à 5 semaines	De bas à moyens	Moyens	Potentiellement excellente	

Suite

Système testant	Durée de l'épreuve	Valeur *		Facilité relative de la détection	
		Frais courants	Frais de mise en exploitation	Mutations géniques	Aberrations chromosomiques
Insectes					
Drosophila melanogaster					
Mutations géniques	2 à 7 semaines	Moyens	Moyens	De bonne à excellente	
Aberrations chromosomiques	2 à 7 semaines	*Idem*	*Idem*		De bonne à excellente
Mammifères					
Létaux dominants	2 à 4 mois	De moyens à élevés	*Idem*		Inconnue
Translocations	5 à 7 mois	*Idem*	*Idem*		Potentiellement très bonne
Analyse cytogénétique des cellules sanguines ou médullaires	1 à 5 semaines	Moyens ***			Potentiellement bonne
Mutations des loci spécifiques	2 à 3 mois	D'élevés à très élevés	D'élevés à très élevés	Inconnue	

* Les frais se mesurant d'après le programme réalisé de l'expérience et de la quantité de substances testées simultanément (en dollars) : très bas, moins de 1000 dollars ; bas, de 1000 à 5000 ; moyens, de 5000 à 10 000 ; élevés, de 10 000 à 20 000 ; très élevés, de 25 000 et plus.

** En ce qui concerne les laboratoires de notre pays, le prix de ce test est de 580 à 735 roubles.

*** En ce qui concerne les laboratoires de notre pays, les investissements d'équipement et de réactifs s'élèvent à 9000 roubles.

L'analyse du matériel actuellement accumulé montre que les effets les plus typiques des différents mutagènes sur les chromosomes sont [29] : l'effet cytostatique, c'est-à-dire l'inhibition du cycle cellulaire ; une spécificité d'action sous forme d'un effet prédominant à telle ou telle phase du cycle cellulaire ; la présence dans certains cas, outre les réactions directes avec la molécule d'ADN, d'un effet mutagène indirect par atteinte des systèmes enzymatiques ; la mutagénèse prolongée dans plusieurs divisions cellulaires suivantes due, sans doute, à ce que les potentialités de mutation sont repliquées dans les processus de synthèse d'ADN ; les ruptures d'une chaîne de l'ADN et la détérioration des bases entraînant des lésions persistantes passant par plusieurs cycles cellulaires.

La méthode la plus courante d'analyse des remaniements structuraux des chromosomes dans les cultures tissulaires, utilisée pour évaluer l'effet des mutagènes, bien qu'elle soit actuellement le mode le plus précis de dénombrer les mutations dans les cellules humaines, présente certains inconvénients.

1° La méthode n'est efficace qu'en présence de fortes doses de mutagènes. Une des particularités des événements génétiques responsables des anomalies chromosomiques est un effet létal prononcé dans les périodes intrautérine et postnatale. Parmi les causes déterminant le fardeau héréditaire que pèse sur la population actuelle, les modifications chromosomiques ne sont identifiées que dans 0,61 % des cas.

2° Malheureusement, il n'existe pas encore de connexion quantitative certaine entre l'aptitude du mutagène à provoquer des cassures chromosomiques et la faculté d'induire des mutations géniques [67].

3° La réalisation de la méthode exige un grand nombre de personnel qualifié.

L'élaboration et le perfectionnement des méthodes cytologiques se poursuivent intensément. Selon certaines données, le pouvoir discriminateur de ces méthodes augmente quand on procède au dénombrement des micronoyaux, formations rejetées dans le cytoplasme par suite d'un remaniement chromosomique non équilibré [63], au calcul des échanges de chromatides sœurs en utilisant un colorant fluorescent permettant une coloration différentielle des chromatides [56, 70].

Dernièrement, on accorde toujours plus d'attention à l'étude expérimentale des possibilités d'enregistrer les mutations géniques par des méthodes biochimiques, en particulier par l'électrophorèse des protéines, du sérum et des hématies sur gel de polyacrylamide ou d'amidon [23, 67]. On peut identifier plus de 100 enzymes par la méthode électrophorétique. La mise en évidence de mutants repose sur un changement de mobilité de la molécule protéique dans le champ électrique. Ainsi, les méthodes biochimiques enregistrent une plus large gamme d'événements de mutation au niveau moléculaire que

toute autre méthode. La cause du changement de charge de la molécule protéique peut être la substitution d'une paire de bases ou l'insertion d'une ou de plusieurs paires de bases dans l'ADN du gène. Toutefois, il faut tenir compte de ce qu'un tiers seulement de toutes les substitutions génétiques possibles dans l'ADN entraîne une variation de charge des protéines.

On considère comme prometteur pour être utilisé en qualité de test de l'activité mutagène et cancérogène l'usage des méthodes biochimiques de détermination de l'activité réparatrice des cellules de mammifères (synthèse non planifiée d'ADN, estimation de la masse moléculaire d'ADN de différents types). De plus, on peut juger des modifications tant par la dégradation d'ADN que par la restauration de sa structure [43, 64].

Malgré l'imperfection de chaque système testant pris en particulier, leur utilisation a apporté de la clarté à plusieurs questions soulevées par la pollution de l'environnement par des mutagènes: les indices quantitatifs d'action mutagène de la radiation ionisante ont été admis; les effets génétiques de l'action complexe de la radiation, des rayons UV et de certains mutagènes chimiques ont été décrits (sensibilisation, effets additif et synergique, phénomènes de défense quand l'influence du mutagène chimique diminue l'efficacité de la radiation); le danger génétique de certains pesticides a été démontré [29].

Il faut considérer comme l'acquis le plus important de l'emploi des systèmes testants la découverte d'une connexion évidente entre l'effet mutagène et l'effet cancérogène de différents composés chimiques. Ceci a été facilité par une large utilisation des systèmes dits d'Ames [6, 7]. En étudiant plus de 300 composés, le groupe d'Ames a établi dans 90 % des cas un lien étroit entre le potentiel cancérogène et le potentiel mutagène des corps chimiques biologiquement actifs [47, 48].

L'emploi des tests d'Ames a permis aux chercheurs japonais de mettre en évidence la haute activité mutagène et cancérogène potentielle du furylfuramide, additif alimentaire largement utilisé au Japon depuis 1965 et considéré jusqu'ici comme inoffensif [35]. Etudiant les causes de la grande fréquence du cancer de l'estomac au Japon, le même groupe de chercheurs a montré qu'au cours de la cuisson de la viande, comme le fait par tradition la population japonaise, certains acides aminés se transforment en métabolites dont l'activité cancérogène dépasse même celle bien connue de l'aflatoxine B (par exemple, les produits de transformation du tryptophane).

Les hydrocarbures aromatiques cycliques (HAC) ont également fait l'objet d'une intense étude génétique [55]. La plupart des composés de cette classe se forment par combustion incomplète des substances organiques à de hautes températures. Un des composés connus de ce groupe (ce qui a permis de le prendre comme indicateur de

groupe) [60] est le benzopyrène déversé dans l'environnement avec les gaz d'échappement, la suie, des fumées diverses. Les récepteurs d'HAC dans la cellule vivante sont l'ADN, dans une moindre mesure l'ARN et les protéines. Les HAC cancérogènes sont mutagènes et provoquent des mutations géniques [6]. La cause de l'effet cancérogène et mutagène est que les hydrocarbures s'intercalent entre les bases de l'ADN et que la formation de liaisons covalentes avec les nucléotides renforce le potentiel cancérogène du composé chimique.

La particularité essentielle de l'influence de nombreux mutagènes chimiques, présents à différentes doses dans l'environnement, sur l'information héréditaire des organismes vivants est le caractère combiné de leur influence, qui s'écarte des lois d'après lesquelles agit chaque mutagène en particulier.

Récemment, un groupe de chercheurs américains a effectué, avec le système testant d'Ames, un travail pour définir l'activité mutagène globale des polluants atmosphériques de 8 régions industrielles urbaines de Californie du Sud, celle de Los Angeles entre autres. Ils utilisaient comme témoin un échantillon d'air prélevé dans la région de Camp Pavik située à plus de 1000 m au-dessus du niveau de la mer dans les montagnes de San Bernandino [57]. Les extraits organiques tirés d'échantillons d'air, recueillis à l'automne 1976 sur des filtres spéciaux en doses différentes (de 0,1 à 50 µl), ont été analysés. Tous les échantillons révélèrent un effet mutagène significatif, augmentant avec la dose sur les systèmes, même sans recourir à l'activation métabolique. Les échantillons prélevés à Los Angeles étaient les plus actifs, les essais de contrôle n'exerçaient pas d'effet mutagène. L'activité cancérogène des échantillons d'air de Los Angeles mise en évidence dès 1954 a été jugée due à une teneur élevée en benzopyrène et autres HAC. Cependant, étant donné que les HAC, le benzopyrène du nombre, ne sont pas actifs en l'absence d'activation métabolique des systèmes testants microbiens, selon les auteurs, les résultats du travail [57] témoignent soit de la présence d'autres mutagènes hautement actifs dans l'air des régions industrielles, soit de l'interaction des HAC avec le smog photochimique typique des villes industrielles de Californie du Sud, aux dépens duquel ils (HAC) produisent un effet mutagène même dans les tests n'utilisant pas l'activation métabolique.

Donc, l'hypothèse énoncée à différentes périodes [10, 18, 68], selon laquelle les mutations affectant les cellules somatiques et la sélection tissulaire sont à l'origine du cancer, est expérimentalement confirmée [15, 65] bien que le lien interne entre la mutagénèse et la cancérogénèse ne soit pas encore élucidé. La confrontation des propriétés mutagènes et cancérogènes des composés est encore rendue plus difficile par le fait qu'une activation de l'organisme par certains métabolites est nécessaire. Dans ces conditions, le précancérogène se transforme en cancérogène, le prémutagène en

mutagène. Probablement que la cause en est la mutagénèse prolongée mise en évidence lors de l'étude de la mutagénèse chimique [20].

La saturation de l'environnement en mutagènes agissant sur l'information génétique de tout ce qui vit et la mise au point des bases moléculaires de la mutagénèse ont posé d'une façon nouvelle les problèmes des normes de quantité de polluants de l'environnement et de leurs doses limites. Cette concentration est considérée comme un seuil. Les normes hygiéniques ou concentrations limites admissibles (CLA) acceptées en U.R.S.S. reposent sur les principes de la présence de seuils biologiques d'action des facteurs chimiques et d'un caractère de nocivité limitant (c'est-à-dire qu'elles reposent sur l'indice le plus sensible). Les CLA sont définies comme la « concentration d'une substance chimique à l'observation de laquelle on ne constate pas d'influence néfaste sur la *santé*, *l'état subjectif et la capacité de travail de la population actuelle et sur les générations humaines futures*, concentration découverte par les méthodes modernes de recherche, et ne faisant pas non plus empirer les conditions de vie hygiéniques de la population » [13]. Cette définition comprend les critères et les principes de l'hygiène soviétique [42] respectant la priorité de la protection de la santé humaine et des indications médicales et la méthodologie de la définition des CLA.

Le progrès des méthodes modernes de recherche — le perfectionnement des systèmes testant l'effet mutagène des polluants, l'étude de l'effet de la mutagénèse prolongée [27] et de la cinétique d'événements de mutation uniques [33] (l'apparition même d'une seule mutation génique ou d'une simple lésion d'un chromosome provoquée par un acte chimique élémentaire ou une seule ionisation peut entraîner les conséquences les plus sérieuses et même un effet létal) —, ainsi que la présence d'effets additifs, sensibilisants et spécifiques, causés par l'action complexe des mutagènes entrant dans la composition de la biosphère, et la présence dans celle-ci de mutagènes dénués d'effet toxique [25] obligent à réviser aussi bien la conception « du seuil » et les normes établies d'après les recherches toxicologiques sur les substances chimiques à effet mutagène que toute la stratégie de la définition des normes. Les calculs montrent [18] qu'une augmentation globale du fond naturel de radiation de 10 rads entraîne l'apparition de 6 millions de porteurs de tares héréditaires dans chaque nouvelle génération. D'autres auteurs [33, 66] considèrent comme inacceptable la conception « du seuil » lors de l'établissement des normes de doses de mutagènes.

La conception du « risque acceptable » [53] suppose que du fait que les seuils sont conventionnels et se rapprochent même de zéro à mesure que les méthodes d'analyse se perfectionnent, l'emploi de termes tels que « niveau de sécurité » et « sécurité de fait » persuade sans aucune justification le public que la notion de sécurité

est un absolu. Par exemple, naguère encore l'affirmation que la question de la possibilité d'établir des normes pour les cancérogènes dans l'environnement est en principe résolue positivement [13] (on peut considérer comme CLA la dose qui provoquera un effet hors des limites de la vie humaine) était vraie. Cependant, la découverte d'une connexion étroite entre la cancérogénèse et la mutagénèse, dont nous avons parlé plus haut, oblige à réviser les CLA établies pour les cancérogènes-mutagènes en raison de leur influence sur la santé des générations futures. Apparemment, il vaut mieux apprécier les mutagènes du point de vue du risque relatif.

Les données accumulées ces derniers temps sur la mutagénèse chimique ont permis à une nouvelle orientation scientifique, la *toxicogénétique* [20], de se constituer. Son but est de prévoir et d'évaluer le dommage potentiel causé au patrimoine génétique collectif des populations par les substances toxiques et par les mutagènes dénués d'effet toxique. Le succès dans ce domaine dépend aussi bien des recherches fondamentales sur la nature moléculaire du processus de mutation, de la création de nouveaux systèmes testants, du perfectionnement de ceux qui existent déjà pour éprouver l'activité mutagène potentielle des composés synthétiques chimiques nouveaux que de l'élaboration d'approches méthodologiques de l'établissement de normes hygiéniques.

Le monitoring génétique des populations

Les systèmes testants élaborés pour déterminer le danger présenté par certains mutagènes parmi les dizaines de milliers qui sont déversés dans l'environnement ne permettent d'évaluer de façon justifiée ni le degré de danger de l'action des polluants sur les populations végétales, animales et humaines, ni la part des fonds, consacrés à la santé publique par n'importe quelle nation, qui doit être dépensée pour soigner les victimes génétiques de la pollution de la biosphère. La solution du problème exige une analyse intégrale de la dynamique du fardeau génétique : l'étude et l'appréciation par la génétique des populations des événements induits par les polluants de la biosphère et frappant tout ce que l'évolution a accumulé au cours de l'histoire (il s'agit des systèmes de polymorphisme génétique équilibré constitués par l'évolution et déterminant le fardeau de ségrégation) [23, 26].

Un important mécanisme tampon contre l'effet immédiat de la pression de mutation est le processus intense du mélange des populations humaines qui se produit [20]. Ainsi, la probabilité de rencontre d'hétérozygotes porteurs de mutations délétères diminue et par là le volume du fardeau génétique latent s'accroît. Les générations futures en feront les frais par le refus de l'homozygote. En même temps, les mutations géniques dominantes et certaines mu-

tations chromosomiques se manifesteront indépendamment de l'augmentation de l'effectif des populations et des degrés de leur mélange. Cette partie émergente, la moins grande, de l'« iceberg » du fardeau génétique peut être estimée dès aujourd'hui.

D'après les données de la littérature mondiale pour 1959-1973 et les matériaux de la statistique internationale [61], les anomalies héréditaires de la population des Caucasoïdes sont estimées à 6 %. D'après les données du Comité scientifique sur les effets de la radiation atomique près l'O.N.U. (United Nations Scientific Commitee on the Effects of Atomic Radiation), ce taux s'est élevé à 10,7 % en 1977. Ainsi se pose le problème d'un monitoring du fond génique des populations, en particulier de la variation du niveau de mutabilité directement d'après les gènes et les chromosomes dans les populations humaines. « Les bases de ce travail doivent être posées de nos jours, il se poursuivra dans l'avenir sur des générations innombrables de l'humanité » [19]. L'élaboration théorique du monitoring des populations humaines soumises à l'action d'un facteur mutagène (radiation ionisante) est devenue importante depuis l'explosion de la bombe atomique à Hiroshima et à Nagasaki.

Trois variantes possibles d'enregistrement des modifications de la structure génétique des populations dans les conditions d'un environnement changeant ont été proposées [50, 52] : le dénombrement des caractéristiques des populations [51] ; la variante de phénotypes « gardiens » ; la variante moléculaire.

Les recherches d'approches du monitoring des séquelles génétiques de la pollution de l'environnement se poursuivent également par l'emploi de méthodes cytogénétiques d'enregistrement des anomalies chromosomiques [51]. Les possibilités d'utilisation de la méthode cytogénétique pour le monitoring des populations, en s'appuyant sur l'expérience accumulée au cours des vingt dernières années, sont exposées dans l'ouvrage [9].

Du fait que les anomalies chromosomiques sont éliminées principalement dans les périodes intrautérine et postnatale précoce, l'usage du monitoring des anomalies chromosomiques ne peut être que la partie constituante d'un programme plus général dans lequel les mutations géniques ont une importance particulière. Sans exercer une influence aussi nocive sur la viabilité de l'organisme que les mutations des chromosomes et des génomes qui n'entrent pas dans le système reproducteur de l'espèce, elles représentent le plus grand danger pour les générations futures. Ainsi faut-il reconnaître l'exceptionnelle difficulté d'un monitoring de la variation du fardeau de mutation chez l'homme. « Dans ce problème, écrit N. Doubinine, un tournant a eu lieu grâce aux recherches de Y. Altoukhov.» La base de la création d'une approche nouvelle du monitoring génétique des populations a été posée par les travaux [2, 5] dont les auteurs ont été les premiers à trouver une voie de délimitation du

fardeau de mutation et de ségrégation dans les populations, ce qui permet d'apprécier l'apport véritable de la pollution de l'environnement à la dynamique du processus de mutation dans les populations. A cette fin, il faut observer un groupe de marqueurs protéiques monomorphes spécialement choisis des gènes.

Les principes théoriques confirmés expérimentalement [23, 24] qui sont à la base de l'approche nouvelle du monitoring génétique des populations ont été formulés dans les travaux [26].

La mise au point des deux orientations ci-dessus décrites dans la génétique actuelle — systèmes testant l'effet mutagène de certains polluants chimiques et monitoring génétique des populations humaines pour apprécier l'accélération possible de la cadence des mutations provoquées par la pollution de l'environnement — amène la nécessité d'organiser un service génétique tant dans certains pays donnés qu'à l'échelle internationale. L'information provenant de ce service peut être à la base d'une planification de mesures médicales, économiques et sociales majeures visant à prévenir les séquelles néfastes de la pollution de l'environnement. L'U.R.S.S. a déjà effectué et effectue un travail scientifique d'organisation aussi bien pour la création de programmes scientifiques que pour l'élaboration de réglements de ce service.

Elaboration de méthodes de protection de l'information génétique contre l'effet mutagène des polluants de la biosphère

Arrêtons-nous brièvement sur une orientation nouvelle, récemment apparue et prometteuse dans l'étude des conséquences génétiques de la pollution de l'environnement: la recherche de moyens pour neutraliser la pression de facteurs mutagènes sur les populations, c'est-à-dire l'étude des phénomènes d'antimutagénèse et du système de réparation des ADN détériorés.

L'étude de l'antimutagénèse élucidant les mécanismes d'action antimutagène de certaines substances permettra de régler et de contrôler les mutations spontanée et provoquée dans les populations. La détection des antimutagènes parmi les composés naturels (les vitamines par exemple) présente un intérêt particulier. L'étude des lois de leur action antimutagène élargira nos connaissances sur le rôle des vitamines dans l'organisme, tandis que l'efficacité des vitamines en faibles concentrations non toxiques sera le gage de leur utilisation pratique contre l'action mutagène de facteurs extrêmes. On a découvert l'universalité de l'activité antimutagène de certaines vitamines: la β-carotine, l'acide ascorbique et surtout l'α-tocophérol. Les antimutagènes agissent à une étape déterminée de la division cellulaire, accélérant une synthèse complémentaire d'ADN. En même temps, le mécanisme de l'activité des antimuta-

gènes est étroitement lié à celui de la défense de la cellule, élaboré au cours de l'évolution et déterminant sa stabilité envers différents agents exogènes, ainsi qu'au complexe de systèmes de réparation d'ADN intracellulaires. Ces derniers assurent la stabilité du génome cellulaire suivant plusieurs paramètres, y compris sa faculté de replication, de recombinaison, de résistance à de nombreux polluants de la biosphère. Les phénomènes de cancérogénèse et de variabilité héréditaire des virus et d'autres organismes sont en liaison étroite avec ces mécanismes cellulaires.

Même si le problème de la pollution de l'environnement était résolu à l'échelle de la planète tout entière, il y aurait toujours des pollutions locales par les produits de certaines branches industrielles. Aussi l'étude des processus de réparation de l'ADN et d'antimutagénèse est-elle nécessaire et d'une grande actualité, elle fait partie du complexe des questions relatives aux séquelles génétiques dues au changement de qualité de l'environnement.

* * *

Les principales orientations examinées plus haut des recherches génétiques dans le domaine de l'étude des séquelles de la saturation du milieu par les polluants mutagènes attestent que la génétique avec sa théorie profondément élaborée de la mutagénèse et de la génétique des populations s'est avérée une des branches des sciences de la Nature les mieux préparées à résoudre les problèmes majeurs de l'existence humaine survenus au cours d'une génération, à associer les recherches fondamentales aux besoins de l'activité économique pratique, tant dans le cadre national qu'à l'échelle mondiale.

La mise au point des principes scientifiques du monitoring des populations et de l'utilisation rationnelle des ressources de la biosphère nous donne la certitude qu'il est possible de prévenir la catastrophe écologique menaçante, et elle nous permettra dans l'avenir d'entreprendre l'organisation de services génétiques nationaux et internationaux contrôlant les phénomènes génétiques dans les populations. Les acquis dans ce domaine de la pensée théorique de la génétique moderne sont appelés à devenir la pierre angulaire d'une nouvelle méthodologie créée de nos jours des formes optimales d'interaction de la société avec son environnement naturel.

RÉFÉRENCES

[1] *Эбботт Дж.* и др. Бюл. Всемир. организ. здравоохр., 1975, 50, 377.
[2] *Алтухов Ю. П. Докл. АН СССР.* 1969, 189, 1115.
[3] *Altukhov Yu. P.* In: XII-th International Conference on Animal Blood Groups and Biochemical Polymorphisms. Budapest, 1970.
[4] *Алтухов Ю. П.* Популяционная генетика рыб. М., Пищевая промышленность, 1974.

[5] *Алтухов Ю. П., Рычков Ю. Г. Ж. общ. Биол.*, 1972, 33, 281.
[6] *Ames B. N. e.a. Proc. Nat. Acad. Sci. USA*, 1972, 69, 3128.
[7] *Ames B. N., Lee F. D., Durston W. E. Ibid.*, 1973, 70, 782.
[8] *Бочков Н. П.* В сб.: Всесторонний анализ окружающей природной среды. Л., Гидрометеоиздат, 1975, 152.
[9] *Бочков Н. П. Цитология и генетика*, 1977, 11, 3, 195.
[10] *Bauer K. H.* Genanderung. Berlin, Springer-Verlag., 1928.
[11] *Bridges B. A. Mutat. Res.*, 1975, 31, 255.
[12] *Brusick D. J., Mager V. Environ. Health Perspect.*, 1973, 6, 83.
[13] *Сидоренко Г. И.* и др. В сб.: Всесторонний анализ окружающей среды. Л., Гидрометеоиздат, 1975, 99.
[14] *Соколов В. Е., Ильенко А. И., Криволуцкий Д. А.* В сб.: Всесторонний анализ окружающей природной среды. Л., Гидрометеоиздат, 1975, 177.
[15] *De Serres F. J. Mutat. Res.*, 1975, 31, 203.
[16] *Добров Г. М.* В кн.: Прогнозирование науки и техники. М., Наука, 1977, 72.
[17] *Drake J. W. Sci.*, 1975, 187, 4176, 503.
[18] *Дубинин Н. П.* Проблемы радиационной генетики. М., Госатомиздат, 1961.
[19] *Дубинин Н. П.* В сб.: Генетические последствия загрязнения окружающей среды. М., ИОГен АН СССР, 1975, 14.
[20] *Дубинин Н. П.* Общая генетика. М., Наука, 1976.
[21] *Дубинин Н. П.* В сб.: Генетические последствия загрязнения окружающей среды. М., Наука, 1977, 3.
[22] *Дубинин Н. П.* и др. В кн.: Успехи современной генетики. 4. М., Наука, 1972, 170.
[23] *Дубинин Н. П.* и др. *Докл. АН СССР*, 1975, 225, 213.
[24] *Дубинин Н. П.* и др. *Докл. АН СССР*, 1976, 230, 967.
[25] *Дубинин Н. П., Пашин Ю. В.* Мутагены окружающей среды. С р. биология. 5. М., Знание, 1977.
[26] *Дубинин Н. П., Алтухов Ю. П.* В сб.: Генетические последствия загрязнения окружающей среды. 2. М., Мысль, 1977, 14.
[27] *Дубинина Л. Г. Генетика*, 1974, 10, 11, 129.
[28] *Дубинина Л. Г.* В сб.: Генетические последствия загрязнения окружающей среды. М., Наука, 1977, с. 89.
[29] *Дубинина Л. Г.* Лейкоциты крови человека — тест-система для оценки мутагенов среды. М., Наука, 1977.
[30] *Дубинина Л. Г., Бигалиев А. Б.* Тест-система оценки мутагенной активности загрязнителей среды в культуре лейкоцитов крови человека. М., 1978.
[31] Evaluation and testicy of drugs for mutagenicity: principles and problems. World Health Organization, Techn. Rept., Ser. 482. Geneva, 1971.
[32] *Федоров Е. К.* Экологический кризис и социальный прогресс. Л., Гидрометеоиздат, 1977.
[33] *Филиппова Л. М.* В сб.: Всесторонний анализ окружающей природной среды. Л., Гидрометеоиздат, 1975, 145.
[34] *Фонштейн Л. М.* и др. Тест-система оценки мутагенной активности загрязнителей среды на *Salmonella*. М., 1977.
[35] *Fox J. L. C and EN*, 1977, Dec. 12, 34.
[36] *Гар К. А.* Инсектициды в сельском хозяйстве. М., Колос, 1974.
[37] *Gabridge M. G. Proc. Soc. Exp. Biol. Med.*, 1969, 130, 831.
[38] Генетические последствия радиации — приложение H (документ A/AC, 82/R 347, 25 января 1977 г., подготовленный Научным комитетом ООН по действию атомной радиации).
[39] *Green S., Springer J. Environ. Health Perspect.*, 1973, 6, 37.

[40] *Касьяненко А. Г., Королева Н. С.* Оценка генетической опасности пестицидов. Пущино, 1977.
[41] *Korte F.* In: Comparative studies of food and environmental contamination. Vienna, 1974, 43.
[42] *Красовский Г. Н., Сутокская И. В.* О критериях определения допустимой нагрузки загрязнения на окружающую среду. Экспресс-информ. ВИНИТИ, 10. М., 1977.
[43] Критерии необходимых и достаточных тест-систем для идентификации потенциальных мутагенных и канцерогенных факторов в окружающей среде. Тезисы докл. сов. участников V советско-американского симпозиума. Баку-Пущино, 1978. М., 1978.
[44] *Куринный А. И., Шликская М. А.* В кн.: Исследование пестицидов как мутагенов внешней среды. Киев, Наукова думка, 1976, 4.
[45] *Малашенко А. М.* В сб.: Генетические последствия загрязнения окружающей среды. 2. М., Мысль, 1977, 69.
[46] *Malling H. V., Frantz C. N. Environ. Health Perspect.*, 1973, 6, 71.
[47] *McCann J.* e.a. *Proc. Nat. Acad. Sci. USA*, 1975, 72, 5135.
[48] *McCann J., Ames B. N. Ibid.*, 1976, 73, 950.
[49] *Мельников Н. Н.* В кн.: Химия и технология пестицидов. М., Химия, 1974, 16.
[50] *Neel J. V. Perspect. Biol. and Med.*, 1971, 14, 522.
[51] *Neel J. V.* Changing perspectives on the genetic effects of radiation. Springfield, 1963, VIII and 97.
[52] *Neel J. V. Mutat. Res.*, 1974, 26, 319.
[53] *Ньюилл В. А.* В сб.: Всесторонний анализ окружающей природной среды. Л., Гидрометеоиздат, 1975, 121.
[54] *Николс У.* В сб.: Генетические последствия загрязнения окружающей среды, М., Наука, 1977, 101.
[55] *Пашин Ю. В., Бахитова Л. М.* В сб.: Генетические последствия загрязнения окружающей среды. 2. М., Мысль, 1977, 46.
[56] *Perry P., Wolff S. Nature*, 1974, 251, 156.
[57] *Pitts J. N.* e.a. *Toxicology Lett.*, 1977, 1, 65.
[58] *Ray V. A., Hyneck M. L. Environ. Health Perspect.*, 1973, 6, 27.
[59] *Russel W. L.* Cold Spring Harbor Sympos. Quantit. Biology., 1951, 16, 327.
[60] *Шабад Л. М.* О циркуляции канцерогенов в окружающей среде. М., Медицина, 1973.
[61] *Шалл Дж.* В сб.: Генетические последствия загрязнения окружающей среды. М., Наука, 1977, 31.
[62] *Шеридан В.* В сб.: Генетические последствия загрязнения окружающей среды. М., Наука, 1977, 95.
[63] *Schmid W. Mutat. Res.* 1975, 31, 9.
[64] *Синельщикова Т. А., Киркова З. С., Засухина Г. Д.* В сб.: Генетические последствия загрязнения окружающей среды. 2. М., Мысль, 1977, 92.
[65] Теоретические и практические подходы к проблеме мутагенеза и канцерогенеза окружающей среды. Материалы советских участников III советско-американского симпозиума. М., Гидрометеоиздат, 1976.
[66] *Тимаков В. Д., Бочков Н. П.* В кн.: Научно-техническая революция и человек. М., Наука, 1977, 128.
[67] *Valcovic L. R., Malling H. V. Environ. Health Perspect.*, 1973, 6, 201.
[68] *Watson J. D.* Molecular biology of the gene. New York, WA Benjamin, 1970.
[69] *Захаров И. А.* В сб.: Генетические последствия загрязнения окружающей среды. 2. М., Мысль, 1977, 65.
[70] *Zatt S. A. Proc. Nat. Acad. Sci. USA*, 1974, 71, 3162.

MONITORING GÉNÉTIQUE DES POPULATIONS HUMAINES EN ACCORD AVEC L'ÉTAT DE L'ENVIRONNEMENT

Y. Altoukhov

De nos jours, dans les conditions de la pollution et d'autres modifications de la biosphère, l'évaluation et la prévision de la dynamique du fardeau génétique en raison de la mutation et de la recombinaison des gènes acquièrent une importance particulière. Bien que certains affirment qu'il est impossible d'estimer la mutabilité changeante dans la population humaine actuelle [24], une approche conduisant à des conclusions opposées [6, 9, 26-31] a été développée ces dernières années.

En nous appuyant sur ces données, nous examinerons les questions suivantes: 1° structure du fardeau héréditaire qui pèse sur la population humaine et principe de la différenciation du fardeau de mutation et de ségrégation; 2° modèle de monitoring génétique des populations en accord avec l'état de l'environnement; 3° justifications du modèle par les faits: identifications des mutations *de novo* et des recombinaisons des gènes lors du screening sélectif de la population.

Structure du fardeau héréditaire de la population humaine. Principe de la différenciation du fardeau de mutation et du fardeau de ségrégation *

D'après les estimations du Comité scientifique sur les effets de la radiation atomique près l'O.N.U. (United Nations Scientific Comitee on the Effects of Atomic Radiation) [1977], le fardeau héréditaire qui pèse sur la population actuelle des Caucasoïdes est de 10,5 %.

Cependant, on observe de nos jours une forte croissance des maladies dites du siècle: cancer et maladie ischémique du cœur,

* Le fardeau de ségrégation est ici traité comme une composante du fardeau héréditaire déterminé par l'apparition dans la population de génotypes pour un ou plusieurs loci polymorphes pernicieux pour l'adaptation.

© Издательство Наука, 1981.

pandémies de grippe, affections aiguës des voies aériennes et autres infections d'origine virale. La fréquence des allergies et des troubles mentaux a augmenté, de nouvelles maladies d'étiologie inconnue sont apparues. En fait, tout ce groupe d'affections, sauf les cancéreuses, n'est pas étudié du point de vue de la génétique, bien qu'il y ait de fortes raisons de croire qu'ici la contribution de l'hérédité soit également très importante [21].

Quand on évalue le volume du fardeau génétique dans la population humaine, il faut tenir compte de ce qu'elle est soumise à la sélection au cours d'une très grande période de l'histoire naturelle, y compris l'époque actuelle. Le cas de l'anémie drépanocytaire minutieusement étudié et largement connu est l'exemple le plus éclatant de l'adaptation à un milieu spécifique basée sur la prédominance sélective des hétérozygotes que « paye » chaque année la mort d'enfants homozygotes pour le gène anomal.

Le « prix » de l'adaptation n'est pas toujours aussi élevé. Cependant, la présence de la composante sélective est prouvée directement ou indirectement pour des systèmes de polymorphisme génétique tels que les groupes sanguins ABO, MN, Rhésus [19, 20, 22, 35, 51, 52, 54, 57, 65, 75, 77]. Mais au cours des 10 à 15 dernières années, grâce aux progrès de la génétique biochimique et, avant tout, des investigations électrophorétiques des protéines en tant que marqueurs des gènes structuraux correspondants, on a découvert une ampleur de la variabilité héréditaire des populations telle qu'il aurait été, naguère encore, impossible de l'imaginer. J. Hubby et R. Lewontin sur la drosophile [43, 50], H. Harris sur l'homme [36, 38], C. Manwell et C. Baker sur les oiseaux [17, 18] ont montré que dans beaucoup de populations naturelles jusqu'à 30 % et plus du nombre des loci étudiés étaient polymorphes et, par conséquent, chaque individu, en moyenne, peut être hétérozygote pour 8 à 15 % de ses gènes. Plus tard, ces découvertes ont été confirmées par les expériences sur d'autres espèces [16, 53, 72, 76, v. également 56].

Le nombre de travaux dans ce domaine croît sans cesse et fait renaître, à un nouveau niveau, la vieille discussion de savoir quel facteur de la dynamique des populations — la dérive génétique ou la sélection — est responsable d'une diversité génétique sans précédent. Nous ne nous posons pas pour tâche d'étudier l'argumentation principale des parties polémisantes, que ceux qui le désirent s'adressent aux travaux correspondants [48, 49].

Signalons seulement que le sens biologique du polymorphisme des protéines ne peut être compris qu'en examinant en même temps les particularités de la structure des populations et du monomorphisme génétique découvert par l'étude électrophorétique des protéines à fonction enzymatique et non enzymatique dans les populations de différentes espèces, y compris l'homme [1, 2, 4, 7, 13-16].

L'approche comparative a montré qu'à côté des protéines poly-

morphes on voit toujours apparaître dans les populations des systèmes protéiques génétiquement invariants, ne ségrégant pas et représentés sur toute l'aire de l'espèce par un seul type prédominant. La gamme de fréquences de telles variantes montre qu'elles ne surviennent que grâce à des mutations itératives.

Mais même si l'on ne subdivise pas les loci contrôlant la synthèse des protéines en polymorphes et monomorphes, tout le pool des allèles dans la population peut être différencié en deux groupes séparés par un hiatus net : 1° les allèles distribués largement et se rencontrant à une grande fréquence ($P > 0{,}01$-$0{,}05$) et 2° les allèles rares, « private * » et se rencontrant à une basse fréquence ($P < 0{,}01$-$0{,}05$) [39, 41]. Plus tard, J. Neel [62, 63, v. également 47, 69] a aussi attiré l'attention sur ce phénomène.

Nous avons postulé [14] que la coexistence simultanée dans la Nature des phénomènes de poly et monomorphisme des protéines en tant que marqueurs géniques peut refléter une dualité de l'organisation fonctionnelle du génome : si le polymorphisme représentant un système souple et mobile d'adaptation de l'espèce à un environnement variant normalement n'est que relativement relié à la viabilité, le monomorphisme est une condition *sine qua non* de l'existence de l'espèce. La lésion par mutation des gènes invariants, spécifiques d'espèce (ou d'une partie du gène) et responsables des fonctions vitales doit entraîner l'élimination des individus ou une forte baisse de leur viabilité.

En effet, depuis longtemps, les chercheurs rattachent les variantes héréditaires rares de protéines identifiées par l'électrophorèse à des troubles du métabolisme. Sur ce plan sont très significatives les annexes au livre de H. Harris [39] ainsi que des données plus nouvelles [42]. Toute une série d'exemples de troubles innés du métabolisme présentés par les auteurs se trouvent liés justement à des protéines dont le polymorphisme génétique n'est pas identifié.

On peut interpréter dans un même sens les résultats de travaux reposant sur l'étude exhaustive de la séquence des acides aminés dans les protéines. Il s'est avéré que la sélection naturelle rejette toutes les mutations affectant les secteurs de la molécule protéique dont dépend la fonction principale du produit génique. Par exemple, chez différentes espèces de vertébrés la lactate-déshydrogénase a une composition différente en acides aminés. Cependant, le centre catalytique de l'enzyme, constitué de 12 radicaux d'acides aminés, reste identique dans tous les cas [45, 71]. Les radicaux d'histidine localisés aux sites fonctionnellement les plus importants — 58, 87, 63 et 92 — des chaînes α et β de la molécule de globine restent tout à fait invariants. Ce monomorphisme se maintient depuis 500 millions d'années de l'histoire naturelle chez de nombreux

* Cité d'après Neel. (*N.d.R.*)

groupes de vertébrés [44]. La substitution de la tyrosine à l'une quelconque de ces histidines provoque une hémoglobinopathie chez l'homme. On peut considérablement élargir le nombre d'exemples de ce genre [v. par exemple 47, 49, 68].

D'un autre côté, l'étude planifiée d'une certaine façon de la génétique des populations naturelles, avant tout en tenant compte de leur structure interne de subpopulations historiquement constituée, a permis d'élucider la contribution des différents facteurs évolutifs au maintien du polymorphisme génétique des protéines. En accord avec les modèles de populations subdivisées [80, 81], on a découvert que la différenciation génétique de l'espèce n'est pas épuisée par l'apport de l'un des facteurs: sélection, migration ou dérive génétique, mais qu'elle est rattachée à une superposition continue de leurs interactions qui ont pour résultat un processus stationnaire [3, 4, 5, 8, 11, 13, 15, 16, 73, 74]. En tout cas, l'analyse des distributions stationnaires des fréquences géniques dans la population subdivisée permet de faire ressortir une composante de la sélection dirigée au profit des génotypes hétérozygotes des systèmes protéiques polymorphes dans la période postnatale [4, 15, 16].

C'est justement en raison de l'importance universelle du polymorphisme des protéines en tant que stratégie adaptative de l'espèce que les mêmes systèmes de polymorphisme s'avèrent communs à des groupes taxonomiques notoirement différents. C'est ainsi que s'établit l'extrême ancienneté de ce genre de variabilité héréditaire homologue qui s'est constituée bien avant la division évolutive de telles ou telles espèces.

Ainsi, les données de la génétique biochimique, d'une part, mettent en évidence la saturation de l'espèce par des allèles anciens du point de vue évolutif et entrés dans son patrimoine génétique intégré en raison de l'interaction continue de la dérive génétique, de la migration et de la sélection et, d'autre part, permettent de détecter des gène jeunes maintenus dans les populations seulement grâce à l'équilibre entre la mutation spontanée (ou induite) et la sélection. C'est précisément l'observation d'un groupe de marqueurs protéiques monomorphes spécialement choisis de gènes ou, plus largement, l'observation de tout le pool d'allèles rares qui permet de saisir une accélération notable de la cadence de la mutation dans une population bisexuée aussi sûrement que si l'on avait pris, comme objet de recherche, des clones ou des lignées pures soumis à l'action des facteurs extérieurs correspondants. Cette conclusion a été faite sur l'exemple des populations gynogénétiques unisexuées de carassin argenté (*Carassius auratus gibelio* Bloch) habitant des bassins spéciaux pollués expérimentalement par des radionucléides [31].

On a découvert une différenciation nette des populations comparées déterminée pour beaucoup par la mutation due aux radio-

Tableau 1

Particularités radioécologiques et génétiques des populations étudiées

Bassin, n°	Concentration des radionucléides, Ci/l · 10^{-11}		Nombre de poissons examinés	Paramètres caractérisant la diversité génétique		
	^{90}Sr	^{137}Cs		Nombre total de variantes électrophorétiques de protéine	Nombre total de variantes nouvelles et rares *	Mesure informationnelle de la diversité, $H \pm m$
1	0,65	0,23	73	43	5	4,8±0,20
2	1,0	0,08	75	36	1	2,9±0,16
3	3	4,0	77	57	19	8,8±0,23
4	2,2	1,5	76	47	15	6,0±0,25
5	230	3,0	152	62	24	9,7±0,18
6	700	2,5	137	56	20	9,0±0,15

* Variantes rares de protéines ($P < 0,05$) ou variantes non constatées chez les populations témoins.

nucléides déversés dans les bassins (*Tableau 1*, fig. 1). Dans l'hypothèse où toute la diversité dans les bassins témoins reflète essentiellement la pression de mutation spontanée, nous avons évalué la contribution de la mutagénèse induite. Des calculs simples ont montré que dans les bassins à niveau intermédiaire et maximal de pollution radioactive, la vitesse de mutation par locus par génération est en moyenne de 10^{-3} à 10^{-2}. Si, compte tenu de données connues [58, 59, 62, 64, 68, 78], on prend pour fréquence de mutation des gènes contrôlant la synthèse des protéines à fonction enzymatique ou non enzymatique 10^{-5}-10^{-6}, ceci signifie une accélération du rythme de la mutation de 2 à 3 ordres de grandeur au moins. De plus, les populations expérimentales diffèrent des témoins par un taux élevé des anomalies morphologiques et par une baisse de la fertilité. Apparemment, l'étude des variantes électrophorétiques rares de protéines dans les populations humaines donne également une clé à l'estimation des cadences de la mutation. Cependant, la réalisation pratique du programme, surtout s'il suppose un screening de masse de la population [61, 63], est liée à de sérieuses difficultés techniques: pour un enregistrement crédible d'une augmentation de 50 % de la vitesse moyenne des mutations, il faut examiner les prises de sang de centaines de milliers et même de millions de sujets pour 20 marqueurs protéiques des gènes [61, 64].

Sans rejeter ce procédé, nous estimons cependant plus justifiée une autre approche dans le cadre de laquelle l'analyse des marqueurs protéiques des gènes fait partie d'un programme plus général de

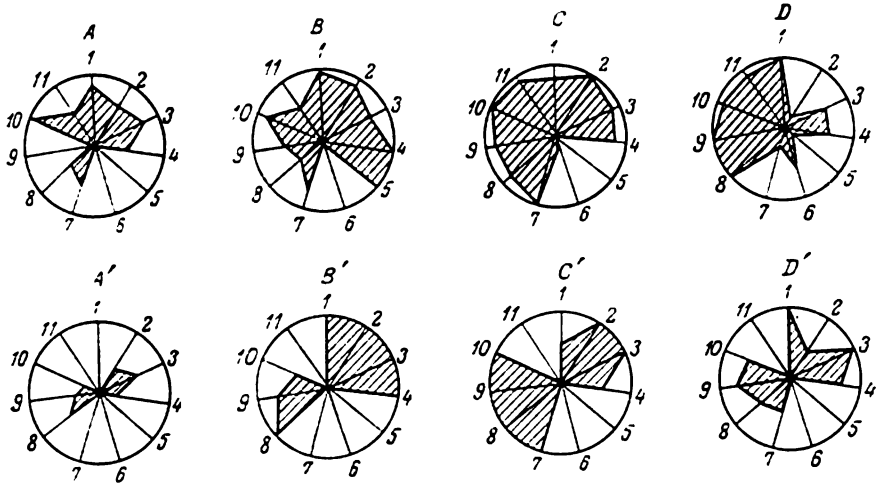


Fig. 1. Polygones combinatoires figurant la diversité génétique des populations étudiées traduite par la mesure informationnelle de Shannon (*A-D*) et évaluée d'après la fréquence des variantes électrophorétiques rares (*A'-D'*):

A, *A'* — les bassins *1* et *2* (globalement); *B*, *B'* — les bassins *3* et *4* (globalement); *C*, *C'* — le bassin *5*; *D*, *D'* — le bassin *6*.

Les données globales sont présentées afin d'égaliser les effectifs des échantillons comparés. On a porté sur les rayons les valeurs des indices correspondants de diversité dans une population donnée, prises en fractions de la valeur maximale de ces indices (100 %) dans n'importe quelle autre population. Zéro se trouve au centre du cercle. On a figuré les protéines suivantes: *1* — alcool-déshydrogénase; *2*, *5* et *8* — différentes zones de l'estérase; *6* — malate-déshydrogénase sérique; *7* — glucose-6-phosphate-déshydrogénase; *9* — catalase; *10* — malate-déshydrogénase musculaire; *11* — myogènes

monitoring biologique. Son examen, compte tenu de la spécificité de la population humaine en tant qu'objet de recherche génétique, est le thème de la partie suivante de l'article.

Les populations humaines en tant qu'objet de recherche génétique suivant l'état de l'environnement. Modèle de monitoring

A la différence du modèle idéal de Hardy, les populations naturelles réelles habitent toujours un milieu concret constitué au cours de l'histoire et y maintiennent un équilibre dynamique dans une longue suite de générations. Telles sont pratiquement les populations de toutes les espèces bisexuées ou, tout au moins, elles étaient telles tant qu'elles ne sont pas devenues l'objet de pressions anthropogènes sortant des limites de l'optimum d'adaptation historiquement constitué [13]. La condition essentielle de la stabilisation des populations est leur organisation systémique: leur subdivision en subpopulations demi-isolées à un effectif génétiquement actif li-

mité. Il est évident que dans l'ensemble de subpopulations à petit effectif interagissant faiblement la valeur du fardeau génétique doit être stable, car les mutations récessives délétères sont inévitablement éliminées au cours de la longue histoire du système.

Apparemment, ce type de structure de populations au caractère spécifique de phénomènes génétiques est propre à la plupart des espèces qui, dans la Nature, se reproduisent par croisement. Cependant, chez l'homme une organisation des populations de ce genre qui lui a été propre durant la plus grande période de son histoire naturelle [Neel et al., 1964] ne se rencontre plus, de nos jours, qu'aux confins de l'Œcumène chez les isolats anciens des petits peuples. Actuellement, la grande majorité de la population a en propre un type de structure fondamentalement différent. La direction et l'intensité de la migration varient de nos jours à un degré tel que l'on passe de l'état de population subdivisée à une large panmixie. On observe une forte augmentation des effectifs général et génétiquement actif, ce qui est particulièrement typique des grandes villes des pays économiquement développés.

Une claire compréhension de ces tendances permet de tracer une approche générale de l'évaluation des phénomènes génétiques dans les populations humaines en raison du problème de la saturation du milieu en mutagènes. Cette approche repose sur les prémisses suivantes.

1° Le milieu d'habitation de l'homme est tellement complexe et tellement hétérogène que même si l'on prête une attention particulière à l'étude des groupes professionnels, il ne sera quand même pas possible d'évaluer l'influence d'un facteur déterminé sur la masse principale de la population : actuellement on voit chaque année se déverser dans la biosphère, en plus de 2 millions déjà enregistrés, des centaines et des milliers de nouveaux composés chimiques. [Evaluation of Genetic Risks of Environmental Chemicals, 1973].

La différenciation du milieu en vue d'une recherche faite par la génétique des populations n'est possible que « de haut en bas » et la première démarche, suggérée par une logique élémentaire, doit être la division du milieu en urbain et rural.

2° La population d'une grande ville est une vaste panmixie. Etant donné la structure actuelle des migrations, elle représente un échantillon adéquat (un noyau) d'une énorme strate anthropologique portant en soi l'information génétique sur l'état du milieu dans une vaste étendue. Par conséquent, une telle population présente un intérêt particulier pour le monitoring et doit être choisie comme objet de recherche.

3° En l'absence de sélection, de mutation ou de migration des gènes dans une grande population panmictique, le processus génétique se déroule suivant le type stationnaire. Le même type de processus est propre à la population en cas d'équilibrement réciproque

de ces facteurs microévolutifs. La prédominance sensible de l'un d'eux est capable de provoquer un processus génétique vectorisé et la dynamique correspondante des paramètres biologiques et morphologiques de la population dans la mesure où ces paramètres sont déterminés par la hérédité.

Comme beaucoup de mutations ont une action néfaste sur la morphogénèse et la physiologie, l'accélération de la cadence de la mutation doit faire son apport spécifique. Des effets dysgéniques analogues peuvent être également provoqués par l'immigration de gènes si se réalisent ainsi de nouvelles combinaisons et associations intra et intergéniques qui abaissent souvent l'adaptabilité [55].

Au niveau des caractères polygéniques, anthropométriques, ce processus génétique défavorable doit se manifester par la « déformation » des distributions correspondantes, d'ordinaire normales, en provoquant leur asymétrie, leur excès, la modification de la dispersion, soit les uns et les autres ensemble.

Ainsi, une partie organique du monitoring génétique doit être le dénombrement de tous les caractères biologiques et morphologiques accessibles définissant la population dans le temps.

4° Néanmoins, les écarts correspondants de la norme en biologie et morphologie de la population peuvent être dus à une action directe du milieu (effets tératogènes) et, par conséquent, une investigation parallèle de la population d'après les gènes mendéliens s'impose. Les possibilités de la génétique immunologique et biochimique permettent d'effectuer cette analyse non pas d'après des loci choisis au hasard, mais sur tout un ensemble de gènes différents par leur localisation et leur fonction. Comme nous l'avons déjà souligné, une attention particulière doit être accordée au groupe de protéines monomorphes.

On est fondé à s'attendre qu'étant donné qu'on utilise un « échantillon » suffisamment représentatif de gènes du génome, les écarts des individus de la norme par les particularités morphologiques polygéniques soient en moyenne corrélés avec le pool de nouveaux gènes issus de la mutation ou de la migration, si par eux-mêmes ou en association avec d'autres gènes ils exercent une influence néfaste sur la viabilité. Deux conséquences importantes s'ensuivent.

Premièrement, la recherche de mutations nouvelles ne s'appuie pas sur le screening de masse de la population mais elle est faite à l'intérieur d'un échantillon non aléatoire d'individus s'écartant de la norme par tout un complexe de caractères.

Deuxièmement, la polémique interminable des « sélectionnistes » et des « neutralistes » au sujet de l'importance biologique du polymorphisme des protéines et des groupes sanguins est laissée de côté. Les principes de l'approche fondés sur la dualité de l'organisation fonctionnelle du génome exigent qu'on cherche des mutations abaissant la viabilité aux stades ontogénétiques précoces, avant tout en

étudiant les nouveau-nés ayant des malformations et les mort-nés (y compris les cas d'avortements spontanés).

5° Avec une telle approche, la question de l'accélération possible des cadences de la mutation chez l'homme pour les gènes, abaissant fortement l'adaptabilité ou absolument liés à la viabilité, est résolue lors de l'estimation de l'écart des distributions correspondantes des caractères polygéniques, déterminé par rapport à un certain niveau « zéro » de lecture pris pour point d'équilibre. Cet intervalle de temps doit être commensurable avec la durée d'existence de la population dans le milieu changeant. Etant donné que le problème de la pollution du milieu par des mutagènes a surgi au cours d'une seule génération, il semble suffisant, en première approximation, de reculer le niveau de référence de 25 à 30 ans. Ceci permet d'embrasser par des observations immédiates trois générations successives simultanément présentes dans la population pour plusieurs caractères, y compris les marqueurs géniques. En principe, ce point de référence peut être reculé encore plus loin dans le temps si l'on prend, pour le justifier, les données de certains collectifs ethniques isolés ayant conservé les principaux traits d'organisation au niveau de la génétique des populations *.

6° Si la pression de mutations est prouvée, l'utilisation des données historiques, démographiques et autres accessibles à l'étude des populations humaines permet de localiser les sources primitives de mutations. En d'autres termes, cette partie du programme suppose la recherche de corrélations entre l'apparition d'une mutation *de novo* et tel ou tel facteur du milieu auquel se sont heurtés (ou se heurtent) les parents de l'enfant.

Plus loin, nous présenterons les justifications de l'approche proposée en nous appuyant sur les résultats de recherches anthropométrique, généticobiochimique et immunologique.

---

* La condition majeure pour comparer les particularités de la mutation dans les populations « natives » et « urbanistes », c'est la communauté d'origine. En l'absence de données de ce genre, des situations très contradictoires sont possibles. Par exemple, les calculs basés sur les fréquences des variantes électrophorétiques de protéines ont montré que les cadences de la mutation chez les Indiens sud-américains sont d'un ou de deux ordres de grandeur plus élevées que parmi les Caucasoïdes [62, 63]. Il est clair que quelle que soit la cause de cet effet — qu'il soit dû à des particularités naturelles de la mutation chez les Indiens, à une contribution de recombinaisons dans les cistrons dont on n'a pas été tenu compte ou à tout autre facteur —, les données obtenues ne valent guère comme point de référence pour, disons, la population d'origine russe de Moscou.

Justifications expérimentales du modèle. Identification du fardeau de ségrégation et de mutation de la population

Nous disposons aujourd'hui de données sur la distribution des phénotypes, des génotypes, des gènes de 500 nouveau-nés et de leurs mères, investigués dans les maternités ordinaires (groupe témoin) et de 200 enfants (nouveau-nés et nourrissons) présentant diverses anomalies morphofonctionnelles de développement, observés dans les cliniques spécialisées.

De nombreux marqueurs polymorphes de gènes gouvernant la synthèse des antigènes érythrocytaires et des diverses protéines sanguines ont été étudiés. En plus des définitions sérologiques et électrophorétiques, on a tenu compte des données généalogiques et démographiques accessibles et l'on a étudié les particularités de la distribution de divers caractères anthropométriques (taille et poids à la naissance, etc.).

Présentons les données anthropométriques pour les nouveau-nés sous forme de distributions des écarts types globaux, par tous les caractères polygéniques, d'un certain phénotype « moyen » pris pour point de référence (fig. 2). Un tel procédé permet de traduire les données obtenues en termes de la conception de la « norme adaptative » et en accord avec l'approche exposée ici qui donne le moyen de suivre l'adaptabilité de la population d'après la variabilité des caractères polygéniques. En effet, on sait depuis longtemps [55] que dans les populations de différentes espèces les mieux adaptés sont justement les individus morphologiquement « moyens ». La même chose a été montrée pour l'homme, tout au moins en ce qui concerne les cas de mortalité accrue parmi les nouveau-nés s'écartant fortement par le poids de la moyenne de la population [46].

En effet, si nous différencions un des groupes d'enfants examinés (nés en 1975 par exemple) en trois sous-groupes de telle sorte que le montre la figure 2 * et que nous les comparions d'après la charge de leurs tares, y compris les malformations, nous constaterions un avantage évident des enfants morphologiquement « moyens » (M^0) par rapport à ceux qui, par l'ensemble des caractères, s'écartent du centre de distribution vers la droite (M^+) et surtout vers la gauche (M^-) (fig. 3). Lorsque l'adaptabilité biologique de la population baisse dans le temps, on doit observer une diminution de la part des phé-

* Etant donné qu'à l'étape actuelle du travail, l'analyse est faite sans tenir compte des exigences de la statistique multidimensionnelle, le procédé proposé ne doit être considéré que comme une première approximation du problème de la création d'une échelle phénotypique unique pour l'identification des individus « moyens ». Néanmoins, l'auteur a étudié l'influence possible de l'« effet de centrage » et s'est convaincu qu'en raison d'une forte corrélation entre les caractères étudiés, son influence sur le résultat final de la classification est négligeable.

notypes « moyens » dans l'échantillon et corrélativement l'augmentation des fréquences des individus ayant des valeurs extrêmes des caractères. Des données de ce genre avaient déjà été examinées auparavant [27, 29].

Maintenant, la tâche consiste à évaluer la mesure dans laquelle la déformation de la courbe est rattachée à une modification de la

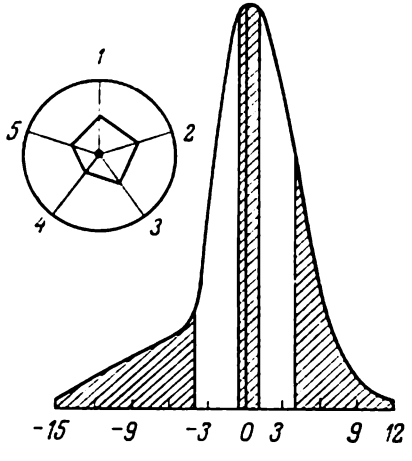


Fig. 2. Distribution des écarts types, par cinq caractères anthropométriques (globalement) des nouveau-nés, d'une moyenne conventionnelle (polygone combinatoire):

Sur l'axe horizontale, $\sum_{1}^{n} \frac{Vi - Mi}{\sigma i}$ où Vi est la valeur individuelle du caractère i dans la population et σ_i l'écart type; sur l'axe vertical, la fréquence des classes

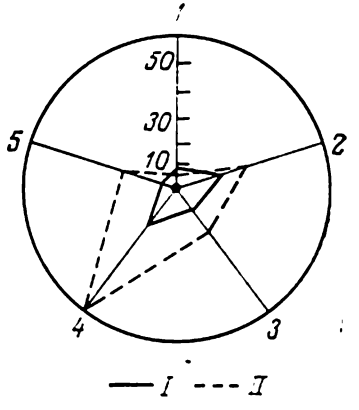


Fig. 3. Niveaux relatifs des écarts morphophysiologiques de la norme découverts dans le sous-groupe de nouveau-nés M^+ (*I*) et M^- (*II*) relativement au groupe M^0 (point au centre du cercle):

Sur le périmètre, la fréquence 0,6; *1* — asphyxie; *2* — cyanose; *3* — ictère; *4* — malformations; *5* — autres anomalies

structure génétique de la population due au fardeau de mutation et de ségrégation. Malgré la différence qualitative entre ces deux composantes principales du fardeau génétique, on constate leur profonde ressemblance: dans ce cas et dans l'autre la population doit être saturée de génotypes rares apportés respectivement par des mutations ou des migrations (recombinaisons) des gènes.

En accord avec le lien postulé plus haut entre les particularités des populations quant aux caractères monofactoriels et polygéniques, ces génotypes rares doivent être plus fréquents justement parmi les individus morphologiquement les plus aberrants.

Examinons à ce sujet les données immunologiques obtenues pour trois sous-groupes de nouveau-nés (v. fig. 2) comparés d'après des caractères morphologiques et pour leurs mères. La figure 4 montre

la distribution empirique des fréquences d'associations génotypiques théoriquement possibles pour les loci des groupes sanguins

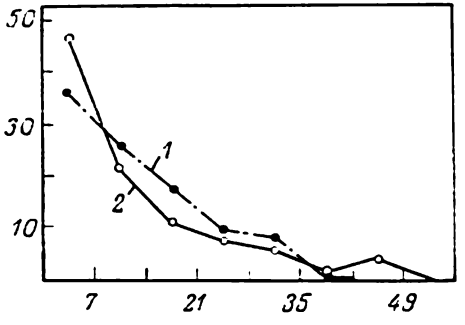


Fig. 4. Distribution des fréquences des associations phénotypiques d'antigènes érythrocytaires des loci ABO, MN, P et Rh parmi les nouveau-nés des groupes M^0 (*1*) et M^- et M^+ globalement (*2*):

Sur l'axe des abscisses, indice de rareté du phénotype par ordre de croissance; sur l'axe des ordonnées, fréquence empirique du phénotype en %

ABO, MN, P et Rh dans les trois sous-groupes de nouveau-nés. On voit clairement que les groupes comparés diffèrent du point de vue

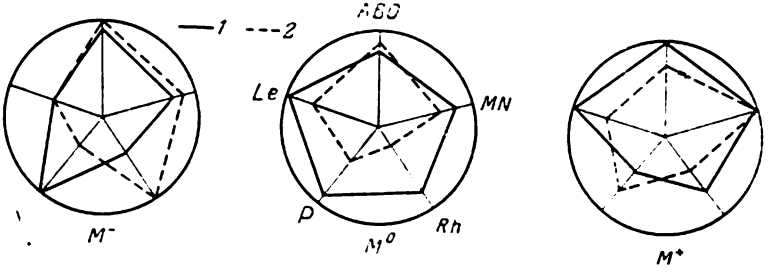


Fig. 5. Polygones combinatoires figurant la diversité génétique pour cinq loci des groupes sanguins de nouveau-nés M^-, M^+, M^0 (*1*) et de leurs mères (*2*) traduite par la valeur de la mesure informationnelle de Shannon $S = -\sum_{i=1}^{n} p_i \ln p_i$ où p_i est la fréquence du phénotype dans le système sanguin correspondant, n le nombre des phénotypes:

Pour les enfants des groupes M^-, M^0, M^+ les valeurs de $S = -\sum_{k=1}^{p} \sum_{i=1}^{n} p_i \ln p_i$ (k est le nombre des loci) sont égales respectivement à 0,96; 1,00; 0,96 et pour leurs mères, à 0,90; 0,82; 0,84

immunogénétique. Si parmi les nouveau-nés morphologiquement « moyens » prédominent les associations de groupes sanguins principalement rencontrées aux fréquences intermédiaires, dans les

Tableau 2

Caractéristique de trois groupes examinés de nouveau-nés d'après les fréquences d'associations antigéniques et certains paramètres démographiques

Groupes examinés	Pourcentage de population urbaine dans deux générations précédentes		Fréquence des mariages entre les individus			Associations antigéniques			
	I	II	des populations géographiquement éloignées	des populations voisines	de la même population	fréquentes (0,03 à 0,12)	moyennes (0,01 à 0,03)	rares (< 0,01)	F_φ
M^0	37,84	55,41	0,2927	0,4634	0,2439	0,3699	0,6164	0,0137	5,79 *
M^-	37,50	69,23	0,3214	0,2500	0,4286	0,4559	0,4853	0,0588	6,82 *
M^+	37,93	70,69	0,4138	0,3103	0,2759	0,4921	0,4603	0,0476	0,33

Remarque. F_φ est le critère de certitude des différences de distributions des associations antigéniques. L'astérisque marque les valeurs de F_φ certaines au niveau $P < 0,01$.

deux autres groupes les plus représentés sont les types antigéniques fréquents et, ce qui est particulièrement important, rares. La différence dans les distributions entre les nouveau-nés M^- et M^+ est insignifiante, mais ces deux groupes se distinguent significativement du groupe M^0 (*Tableau 2*). De même, la diversité génétique figurée par la valeur de la mesure informationnelle de Shannon (fig. 5) s'avère maximale pour les nouveau-nés morphologiquement « moyens ». Par le caractère en question, on découvre des différences évidentes entre les mères et les enfants dans les groupes correspondants: les mères des enfants du groupe M^0 qui ont fourni la descendance la plus diversifiée possèdent la plus faible diversité génétique. Les mères des enfants M^- se caractérisent par un niveau élevé de diversité génétique quantitativement proche de l'indice correspondant de leurs enfants, mais avec une redistribution qualitative de cette information dans la génération suivante. Comme cette redistribution touche aussi le système Rhésus, elle peut être liée à une incompatibilité fœto-maternelle responsable de la naissance d'enfants présentant des signes de malformations. Ces enfants constituent la part prédominante du groupe M^-. La redistribution de l'information génétique est également observée, bien que moins prononcée, dans le groupe M^+.

Par conséquent, la cause principale de si fortes différences dans la structure génétique des groupes de nouveau-nés correspondants se trouve élucidée. Il est évident que la nouvelle information héréditaire découverte dans les groupes M^+ et M^- ne pouvait avoir été apportée que par les pères, c'est-à-dire par migration génique. Elle s'est trouvée répartie inégalement dans les sous-groupes étudiés, les effets défavorables étant surtout accusés parmi les nouveau-nés du groupe M^-. C'est le fardeau de ségrégation de la population et, si notre recherche tenait également compte des variantes électrophorétiques rares de protéines, nous aurions plus de chance de déceler le fardeau de mutation. En effet, les toutes premières analyses électrophorétiques des protéines sanguines d'enfants s'écartant fortement de la moyenne de la population par le poids à la naissance (50 prématurés et 177 portant de multiples malformations grossières) ont montré les différences entre ces enfants et le groupe témoin des nouveau-nés normaux d'après la fréquence des génotypes rares (*Tableau 3*): si l'on n'a découvert dans le groupe témoin que 5 variantes avec mobilité ou activité électrophorétique modifiée, parmi les prématurés et les enfants mal conformés ce nombre était de 15. Ceci signifie que, dans le dernier cas, la fréquence moyenne de gènes rares était environ d'un ordre de grandeur plus élevée que dans le premier [7, 28].

Des données de ce genre ont une importance essentielle. Elles confirment la conjecture, faite plus tôt, sur le lien négatif entre les mutations des gènes monomorphes et la viabilité de l'organisme et, par conséquent, elles autorisent à affirmer que la part prédominante des allèles rares enregistrés par électrophorèse est représentée par des mutations « jeunes » du point de vue évolutif et n'est pas héritée des générations précédentes.

En effet, l'analyse du sang des parents d'un enfant prématuré ayant un génotype variant pour le locus LDH-A (fig. 6) a montré nettement que cette variante manque chez eux, alors que la paternité n'est pas exclue pour 16 loci polymorphes (*Tableau 4*).

On comprend qu'il soit très difficile d'obtenir des prises de sang des parents d'un enfant ayant de multiples anomalies de développement. Néanmoins, nous avons pu examiner 4 familles de ce genre et nous rendre compte qu'ici également le même tableau est observé: les variantes d'estérase érythrocytaire (2 cas), de phosphoglucomutase érythrocytaire (1 cas) (v. fig. 6) et d'albumine sérique (1 cas) découvertes chez les enfants manquent chez les deux progéniteurs. De même que dans le cas de la mutation pour la lactate-déshydrogénase, la paternité n'était pas exclue pour de nombreux gènes polymorphes (v. *Tableau 4*) et, par conséquent, il y a de fortes raisons de conclure que les mutations que nous avons identifiées étaient survenues *de novo*, c'est-à-dire dans une génération donnée. Le fait que les variantes électrophorétiques rares de protéines monomor-

Tableau 3

Fréquence des variantes électrophorétiques rares de protéines chez les nouveau-nés sains, les prématurés et les enfants ayant des malformations

Protéines étudiées par électrophorèse	Nouveau-nés sains		Prématurés et les enfants ayant des malformations	
	Effectif de l'échantillon	Nombre de sujets à variante rare	Effectif de l'échantillon	Nombre de sujets à variante rare
Lactate-déshydrogénase érythrocytaire	498	1 *	227	1
Malate-déshydrogénase érythrocytaire	498	1	227	0
Phosphogluconate-déshydrogénase érythrocytaire	504	1 *	226	2 (1 *)
Estérase érythrocytaire	498	1 *	227	4 (1 *)
Phosphoglucomutase érythrocytaire	427	1	193	3
Phosphatase acide érythrocytaire	298	0	143	1
Superoxyde dismutase érythrocytaire	207	0	133	0
Glutamopyruvate-transaminase érythrocytaire	156	0	82	0
Hémoglobine	498	0	227	2
Transferrine (sidérophiline) sérique	504	0	222	1
Composante spécifique de groupe du sérum	499	0	223	0
Albumine sérique	499	0	227	1
Fréquence moyenne par locus par individu	0,000 312±0,000 140		0,002 035±0,000 525	

* Variantes découvertes chez le même individu.

phes sont corrélées à de graves troubles, pratiquement incompatibles avec la vie, de la morphogénèse rend cette conclusion irréfutable.

Jusqu'à présent, même en étudiant de grands groupes de la population, on n'avait pas découvert de mutations *de novo* de gènes gouvernant la synthèse des enzymes. H. Harris et coauteurs [41] ont généralisé les résultats de longues années de recherche sur des Européens dans le sang desquels on avait décelé des variantes électrophorétiques rares de protéines et montré que dans aucune des 77 familles étudiées la variante rare ne pouvait être attribuée à une mutation dans une génération donnée. Dans tous les cas, la même variante a été trouvée soit chez la mère, soit chez le père, c'est-à-dire qu'elle était héritée de la génération précédente.

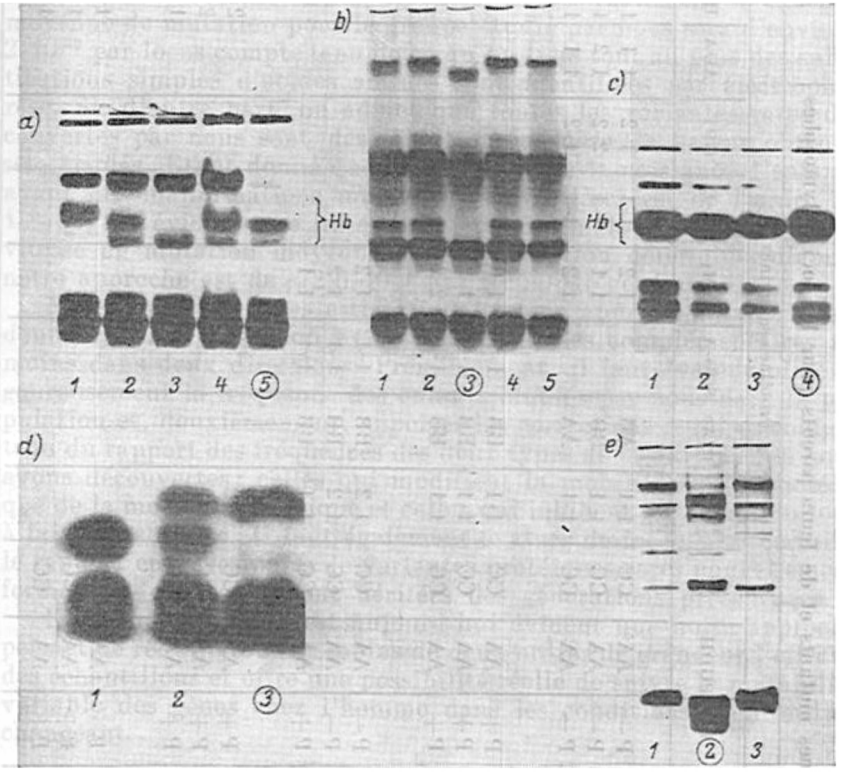


Fig. 6. Variantes électrophorétiques rares de protéines découvertes chez les enfants ayant des malformations (chiffres encerclés):

Electrophorèse sur gel de polyacrylamide, solution tampon THAM-EDTA-borate, pH 8,6. Sur les clichés la cathode est en haut; l'anode, en bas; Hb — hémoglobine; *a* et *c* — estérase érythrocytaire; *b* — lactate-déshydrogénase érythrocytaire; *d* — phosphoglucomutase (sur gel d'amidon); *e* — albumine sérique

Les auteurs ont pu donner une estimation probabiliste de la limite supérieure de la vitesse de mutation pour les gènes gouvernant la synthèse des protéines sériques et érythrocytaires d'après la formule $0{,}05 = (1 - \mu)^n$ où μ est la vitesse de mutation et n le nombre d'allèles découverts par screening. Il s'est avéré que la vitesse de mutation supérieure à $2{,}24 \cdot 10^{-5}$ par gène par génération doit être exclue avec une probabilité $P \geqslant 0{,}95$. D'après une estimation récente de M. Nei [69] profitant des données de Neel [62], la vitesse de mutation pour les gènes contrôlant la synthèse des protéines sanguines est $\sim 7{,}2 \cdot 10^{-6}$.

Si l'on admet que dans notre recherche 5 seulement du nombre des variantes rares de protéines présentées au *Tableau 2* ont pour

Tableau 4

Génotypes des enfants porteurs de gènes mutants et de leurs parents pour les loci polymorphes

Familles investiguées dans lesquelles on a découvert une mutation	Antigènes érythrocytaires				Protéines sanguines									
	ABO	MNSs	Rhésus	P	PGDH	Tf	Gc	P_1Ac	SOD-A	PGM_1	Est-D	GPT	ADA	Hp
Enfant (μLDH-A)	B	MNss	ccDEe	P—	AA	CC	1-1	AB		1-2	1-1	1-2	1-1	1-1
Mère	B	MMSs	ccDEe	P—	AA	CC	1-1	BB		1-1	1-1	2-2	1-2	1-1
Père	B	MNss	ccDEe	P—	AA	CC	1-1	AA		1-2	1-1	1-2	1-1	1-2
Enfant (μPGM_1)	B	MNSs	ccDEe	P+	AA	CC	1-1	BB	1-1	1-0				
Mère	AB	MNSs	ccDEe	P+	AA	CC	1-1	BB	1-1	1-2				
Père	A	MNSs	CcDEe	P+	AA	CC	1-1	BB	1-1	2-2				
Enfant (μEst)	A	MNss	CcDEe	P—	AA	CC	1-2	BB	1-1	1-2				
Mère	A	MNss	CcDee	P—	AA	CC	1-2	BB	1-1	1-2				
Père	OO	MNss	CcDee	P+	AA	CC	1-1	BB	1-1	1-1				
Enfant (μEst)	OO	MM	Rh+	P+	AA	CC	1-2							2-2
Mère	OO	MM	Rh+	P+	AA	CC	1-2							2-2
Père	OO	MM	Rh+	P+	AA	CC	1-2							1-2
Enfant (μAlb)	A	MN	Rh+	P—	AA	CC								1-2
Mère	OO	MN	Rh+	P+	AA	CC								1-2
Père	A	MN	Rh+	P+	AA	CC								1-2

source des mutations « toutes fraîches », alors également la cadence moyenne de mutation pour le groupe étudié par nous sera d'environ $2 \cdot 10^{-3}$ par locus compte tenu de ce qu'un tiers tout au plus des substitutions simples d'acides aminés sont identifiées par électrophorèse. Si, d'autre part, on admet que toutes les variantes rares découvertes par nous sont des mutations *de novo*, la valeur obtenue sera triplée. Etant donné que la fréquence de la naissance d'enfants ayant des malformations multiples et grossières est de l'ordre de 1 %, il est évident que les estimations minimale et optimale de la vitesse de mutation moyenne de la population déterminée d'après notre approche est de $\sim 2\text{-}6 \cdot 10^{-5}$ par locus par génération.

Nous considérons ces estimations comme approximatives, étant donné que leur précision exige des recherches complémentaires au moins dans deux directions. Premièrement, il faut évaleur plus rigoureusement la fréquence des enfants étudiés par nous dans la population et, deuxièmement, apporter les corrections requises compte tenu du rapport des fréquences des deux types de mutations que nous avons découvertes: celles qui modifient la mobilité électrophorétique de la molécule protéique et celles qui inhibent fortement ou tout à fait son activité. Il faut également évaluer de façon plus certaine le rapport entre les parts de variantes protéiques rares nouvellement formées et celles qui sont héritées des générations précédentes.

En même temps, il est aujourd'hui évident que notre approche permet de réduire tout au moins de deux ordres de grandeur l'effectif des échantillons et offre une possibilité réelle de suivre la mutabilité variable des gènes chez l'homme dans les conditions d'un milieu changeant.

CONCLUSION

Dans les parties précédentes de l'article, nous avons justifié l'approche de l'étude du processus génétique défavorable dans les populations humaines et montré l'effet ainsi découvert de l'apparition de génotypes rares, provoqué par l'immigration et la mutation des gènes. Afin de localiser les sources de ce phénomène, il faut, nous y avons insisté, recourir aux données démographiques et généalogiques. Une analyse de ce genre n'a été faite aujourd'hui par nous que pour la composante de ségrégation du fardeau génétique (v. *Tableau 2*). Ces données montrent que dans les généalogies des enfants M^+ et M^- sont plus souvent constatés les mariages entre individus de populations géographiquement éloignées (M^+) ou entre individus de la même population (dans les limites de la même ville ou du même village, M^-), alors que dans le groupe M^0 presque la moitié des mariages se font entre individus de populations voisines dans les limites d'un système historiquement constitué (partie centrale du territoire européen de l'U.R.S.S.). On observe en même temps une

redistribution des parts de la population urbaine et rurale: dans la génération des mères le pourcentage des habitantes des villes a fortement augmenté, ce qui est particulièrement évident dans les groupes M^+ et M^-.

On peut donc tirer la conclusion de l'importance adaptative de la différenciation découverte et la rattacher à un certain optimum d'hétérozygotie. Les données envisagées, tant démographiques que reflétant l'état de santé dans les trois groupes comparés, attestent que cet optimum doit coïncider avec le groupe moyen (forme stabilisante de la sélection).

Par conséquent, la possibilité d'une interprétation génétique plus rigoureuse des phénomènes biologiques dans la population panmictique actuelle s'offre. Il faut y rapporter en premier lieu le phénomène d'accélération et le refus de génotypes moins favorables complétant la part du patrimoine génétique de la population découverte au début des années trente et appelée par la suite *fardeau génétique* [23, 33, v. également 25].

Il devient évident que de nos jours, les causes et les sources de l'accroissement du fardeau génétique dans les populations humaines doivent être cherchées dans la structure des migrations de la population et dans l'état de l'environnement. L'influence des mutations *de novo* des gènes sur l'ampleur des affections des populations était, apparemment, de peu d'importance en comparaison du fardeau des génotypes moins adaptés et refusés dans chaque génération en raison de la présence chez l'homme de complexes de gènes polymorphes en équilibre avec l'environnement concret.

Dans les conditions où la biosphère est saturée de mutagènes, la situation peut brusquement changer et, par conséquent, empirer encore. Il est évident qu'aujourd'hui, l'humanité se heurte à un problème d'une ampleur et de conséquences sociales sans précédent dans toute l'histoire connue. Nous devons faire tout le nécessaire pour obtenir le plus vite possible les estimations quantitatives et réaliser un système de mesures permettant de surmonter les effets indésirables. L'auteur espère que l'approche présentée ici peut jouer un certain rôle dans la solution de ce problème.

RÉFÉRENCES

[1] *Алтухов Ю. П.* О соотношении моно- и полиморфизма гемоглобинов в микроэволюции рыб. *ДАН СССР*, 1969, 189, 1115-1117.

[2] *Altukhov Yu. P.* Biochemical polymorphism in fish and the problem of intraspecific differentiation. In: XII Intern. Conf. Animal. Blood Groups and Biochemical Polymorphism. Budapest, 1970.

[3] *Altukhov Yu. P.* Rational management of marine biological resources in the light of population genetics. In: Propagation marine resources Pacific Ocean. Tokyo, Tokai Univ., 1973, 31-39.

[4] *Алтухов Ю. П.* Популяционная генетика рыб. М., Пищевая промышленность, 1974, 247 с.

[5] *Алтухов Ю. П.* Генетика природных популяций и ресурсы биосферы. *Вестн. АН СССР*, 1975, 10, 37-45.

[6] *Алтухов Ю. П.* Подход к дифференциации мутационного и сегрегационного груза в популяции человека. В кн.: Критерии необходимости и достаточности тест-систем для идентификации мутагенных и канцерогенных факторов окружающей среды. В кн.: Тез. докл. V Сов.-амер. симпоз., М., 1978, 4-6.

[7] *Altukhov Yu. P.* Biochemical population genetics and speciation. *Evolution*, 1982, 36, 6, 1168-1181.

[8] *Алтухов Ю. П., Бернашевская А. Г.* Экспериментальное моделирование динамики генных частот в системе полуизолированных популяций. *ДАН СССР*, 1978, 238, 712-714.

[9] *Altukhov Yu. P., Dubinin N. P., Khilchevskaya R. I.* Genetic monitoring of human populations. In: Abstr. V Intern. Congr. Human Genet. Mexico, 1976, 176-177.

[10] *Алтухов Ю. П., Калабушкин Б. А.* Стабильный полиморфизм в современной и ископаемой популяциях моллюска *Littorina squalida*. *ДАН СССР*, 1974, 215, 1477-1480.

[11] *Алтухов Ю. П., Победоносцева Е. Ю.* Экспериментальное моделирование генетических процессов в подразделенных популяциях. *ДАН СССР*, 1978, 238, 466-469.

[12] *Алтухов Ю. П., Пудовкин А. И., Салменкова Е. А., Коновалов С. М.* Стационарность распределений частот генов лактат-дегидрогеназы и фосфоглюкомутазы в системе субпопуляций локального стада рыб (на примере *Oncorhynchus nerka* Walb.). Сообщение II. Случайный дрейф генов, миграция и отбор как факторы стационарности. *Генетика*, 1975, 11, 54-62.

[13] *Алтухов Ю. П., Рычков Ю. Г.* Популяционные системы и их структурные компоненты. Генетическая стабильность и изменчивость. *Журн. общ. Биол.*, 1970, 31, 507-526.

[14] *Алтухов Ю. П., Рычков Ю. Г.* Генетический мономорфизм видов и его возможное биологическое значение. *Журн. общ. Биол.*, 1972, 33, 281-300.

[15] *Алтухов Ю. П., Салменкова Е. А., Коновалов С. М., Пудовкин А. И.* Стационарность распределений частот генов лактатдегидрогеназы и фосфоглюкомутазы в системе субпопуляций локального стада рыб (на примере *Oncorhynchus nerka* Walb.). Сообщение I. Стабильность стада в поколениях при одновременной изменчивости слагающих их структуру субпопуляций. *Генетика*, 1975, 11, 44-53.

[16] *Алтухов Ю. П., Салменкова Е. А., Омельченко В. Г., Сачко Г. Д., Слынько В. И.* О числе мономорфных и полиморфных локусов в популяции кеты — одного из тетраплоидных видов тихоокеанских лососей. *Генетика*, 1972, 8, 251-259.

[17] *Baker C. M. A., Manwell C.* Molecular genetics of avian proteins. VIII. Eggwhite proteins of the migratory quail *Coturnix coturnix*: new concepts of « hybrid vigour ». *Comp. and Biochem. Physiol.*, 1967, 23, 21-42.

[18] *Baker C. M. A., Manwell C., Labisky R. F., Harper I. A.* Molecular genetics of avian proteins. V. Eggs, blood and tissue proteins of the ringnecked pheasant *Phasianus colchicus*. L. *Comp. and Biochem. Physiol.*, 1966, 17, 467-499.

[19] *Brues A. M.* Selection and polymorphism in the ABO blood groups. *Amer. J. Phys. Anthropol.*, 1954, 12, 559-597.

[20] *Brues A. M.* Stochastic tests of selection in the ABO blood groups. *Amer. J. Phys. Anthropol.*, 1963, 21, p. 287-299.

[21] *Cavalli-Sforza L. L., Bodmer W. F.* The Genetics of human populations. San Francisco, Freeman, 1971, 965 p.

[22] *Chung C. S., Morton N. E.* Selection at the ABO locus. *Amer. J. Hum Genet.*, 1961, 13, 9-27.

[23] *Crow J. F.* Some possibilities for measuring selection intensities in man. In: Natural Selection in Man/Ed. *I. N. Spuhler.* Detroit, 1958, 1-13.
[24] *Crow J. F.* Human population monitoring. In: Chemical mutagens. N.Y., 1971, 2, 591-605.
[25] *Dobzhansky Th.* Introduction for lecture by Professor N. P. Dubinin. In: Proc. XII Intern. Congr. Genetics, Tokyo, 1968, 3, 77.
[26] *Дубинин Н. П., Алтухов Ю. П.* Генетические и биологические процессы в популяциях как индикатор состояния окружающей среды. В кн.: Теоретические и практические подходы к проблеме мутагенности и канцерогенности окружающей среды. М., Гидрометеоиздат, 1976, 38-40.
[27] *Дубинин Н. П., Алтухов Ю. П.* Мониторинг генетических последствий загрязнения окружающей среды: общий подход. В кн.: Генетические последствия загрязнения окружающей среды. М., Мысль, 1977, 14-45.
[28] *Dubinin N. P.* and *Altukhov Yu. P.* Gene mutations (*de novo*) found in electrophoretic studies of blood proteins of infants with anomalous development. *Proc. Nat. Acad. Sci. USA*, 1979, **76**, 5226-5229.
[29] *Дубинин Н. П., Алтухов Ю. П.* Окружающая среда и генетический груз популяций. В сб.: Успехи современной генетики. **10**, Москва, Наука, 1982, 3-26.
[30] *Дубинин Н. П., Алтухов Ю. П., Курбатова О. Л., Сусков И. И.* Интегральная генетическая характеристика «адаптивной нормы» в популяции человека. *ДАН СССР*, 1976, **230**, 957-960.
[31] *Дубинин Н. П., Алтухов Ю. П., Салменкова Е. А., Милишников А. Н., Новикова Т. А.* Анализ мономорфных маркеров генов в популяциях как метод оценки мутагенности среды. *ДАН СССР*, 1975, **225**, 213-216.
[32] *Дубинин Н. П., Алтухов Ю. П., Сусков И. И., Хильчевская Р. И., Афанасьев К. И., Малинина Т. Н., Большакова Н. К., Шурхал А. В., Курбатова О. Л., Прохоровская В. Д.* Экспериментальное обоснование принципов мониторинга генных мутаций у человека. *ДАН СССР*, 1978, **243**, 1313-1316.
[33] *Дубинин Н. П., Гептнер М. А., Никоро З. М., Бессмертная С. Я., Беляева В. Н., Демидова З. А., Кроткова А. П., Постникова Е. Д.* Экспериментальное изучение экогенотипов у *Drosophila melanogaster*. *Биол. Ж.*, 1934, 3, 166-216.
[34] Evaluation of Genetic Risks of Environmental Chemicals. *AMBIO, Special Report*, 1973, 3, 9.
[35] *Glass B.* The action of selection on the principal Rh alleles. *Amer. J. Hum. Genet.*, 1950, 2, 269-278.
[36] *Harris H.* Enzyme polymorphism in man. *Proc. Roy. Soc.* London, 1966, 164, 298-310.
[37] *Harris H.* Genes and enzymes. *Proc. Roy. Soc.* London, 1969, 174, 1-15.
[38] *Harris H.* Enzymes and protein polymorphism in human populations. *Brit. Med. Bull.*, 1969b, 25, 5-13.
[39] *Harris H.* The principles of human biochemical genetics. Amsterdam, North-Holland Publ. Co., 1977, 173 p.
[40] *Harris H., Hopkinson D. A.* Handbook of enzyme electrophoresis in human genetics. Amsterdam, North-Holland Publ. Co., 1978.
[41] *Harris H., Hopkinson D. A., Robson E. B.* The incidence of rare alleles determining electrophoretic variants: data on 43 enzyme loci in man. *Ann. Hum. Genet.*, 1974, 37, 237-253.
[42] *Harris H., Hopkinson D. A.* Handbook of enzyme electrophoresis in human genetics. Amsterdam, North-Holland Publ. Co, 1976.
[43] *Hubby J. L., Lewontin R. C.* A molecular approach to the study of genic heterozygoty in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*. *Genetics*, 1966, 54, 577-594.
[44] *Jukes T. H.* Comparisons of polypeptide chains of globins. *J. Mol. Evol.*, 1971, 1, 46-62.

[45] *Kaplan N. O.* Evolution of dehydrogenases. In: Evolving genes and proteins. N.Y., Acad. Press, 1965, 245-279.
[46] *Karn M. N., Penrose L. S.* Birth weight and gestation time in relation to maternal age, parity and infant survival. *Ann. Eugen.*, 1951, 15, 206.
[47] *Kimura M., Ohta T.* Mutation and evolution at the molecular level. *Genetics*, 1973, 73, 19-35.
[48] *Lewontin R. C.* Population genetics. *Ann Rev. Genet.*, 1973, 7, 1-27.
[49] *Lewontin R. C.* Genetic basis of evolutionary changes. N.Y.-L., Columbia Univ. Press, 1974.
[50] *Lewontin R. C., Hubby J. L.* A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics*, 1966, 54, 595-609.
[51] *Lewontin R. C., Krakauer J.* Distribution of gene frequences as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, 1973, 74, 175-195.
[52] *Livingstone F. B.* Natural selection and ongoing human evolution as illustrated by the ABO blood groups. *Human Biol.*, 1960, 32, 17-27.
[53] *Manwell C., Baker A.* Molecular biology and the origin of species: heterosis, protein polymorphism and animal breeding. L., 1970, 394.
[54] *Matsunaga E., Hirazumi C.* Prezygotic selection in ABO blood groups. *Science*, 1962, 135, 432-434.
[55] *Mayr E.* Animal species and evolution. Harvard Univ. Press, 1965.
[56] Molecular Evolution/Ed. F.J. Ayala. Sunderland. 1976, 277 p.
[57] *Morton N. F., Chung C. S.* Are the MN blood groups maintained by selection? *Amer. J. Hum. Genet.*, 1959, 11, 237-251.
[58] *Mukai T., Cockerham C. C.* Spontaneous isozymes genes in *Drosophila melanogaster*. *Dros. Inform. Serv.*, 1970, 45, 99.
[59] *Mukai T., Cockerham C. C.* Spontaneous mutation rates at enzyme loci in *Drosophila melanogaster*. *Proc. Nat. Acad. Sci. USA*, 1977, 74, 2414-2517.
[60] *Muller H. J.* Our load of mutations. *Amer. J. Hum. Genet.*, 1950, 2, 111-176.
[61] *Neel J. V.* The detection of increased mutation rates in human populations. In: Mutagenic effects of environmental contaminants. N.Y., Acad. Press, 1972, 99-119.
[62] *Neel J. V.* « Private » genetic variants and the frequency of mutation among South American Indians. *Proc. Nat. Acad. Sci. USA*, 1973, 70, 3311-3315.
[63] *Neel J. V.* Development in monitoring human populations for mutation rates. *Mutat. Res.*, 1974, 26, 319-328.
[64] *Neel J. V.* Some considerations pertinent to monitoring human populations for changing mutation rates. In: Well-being of mankind and genetics. Proc. XIV. Internat. Congr. Genet., 1, 1, Moskow, Mir Publ., 1980, 225-238.
[65] *Neel J. V., Schull W. J.* Human heredity. Univ. Chicago Press, 1954.
[66] *Neel J. V., Salzano F. M., Junqueira P. G., Keiter F., Maybury-Lewis D.* Studies on the Xavante Indians of the Brazilian Mato Grosso. *Amer. J. Hum. Genet.*, 1964, 16, 52-140.
[67] *Neel J. V., Salzano F.* Further studies on the Xavante Indians. X. Some hypotheses-generalizations resulting from these studies. *Amer. J. Hum. Genet.*, 1966, 19, 554-574.
[68] *Nei M.* Molecular population genetics and evolution. Amsterdam, North-Holland Publ. Co., 1975.
[69] *Nei M.* Estimation of mutation rates from rare protein variants. *Amer. J. Hum. Genet.*, 1977, 29, 225-232.
[70] *Ohno S.* The enormous diversity in genome size of fish as a reflection of nature's extensive experiments with gene duplications. *Trans. Amer. Fish Soc.*, 1970a, 99, 120-131.

[71] *Ohno S.* Evolution by gene duplication. Springer-Verl., 1970b.
[72] *Petras M. L., Reimer J. D., Biddl F. G., Martin J. E., Linton R. S.* Studies of natural populations of *Mus.* V. A survey of nine loci for polymorphisms. *Canad. J. Cytol.*, 1969, 11, 497-513.
[73] *Рычков Ю. Г.* Некоторые популяционно-генетические подходы к антропологии Сибири. *Вопр. антропол.*, 1969, 33, 16-33.
[74] *Рычков Ю. Г., Мовсесян А. А.* Генетико-антропологический анализ распределения аномалий черепа у монголоидов Сибири в связи с проблемой их происхождения. *Тр. Моск. о-ва испытателей природы*, 1972, 43, 114-132.
[75] *Рычков Ю. Г., Русакова О. Л., Раппопорт М. П., Шереметьева В. А.* Факторы генетической дифференциации популяционной системы коренного населения Северной Азии. I. Определение интегрирующих значений адаптивной ценности генотипов группы крови АВО в коренном населении Сибири. *Генетика*, 1973, 9, 136-145.
[76] *Selander R. K., Yang S. Y., Lewontin R. C., Jonson W. E.* Genetic variation in the Horseshoe Crab (*Limulus polyphemus*), a phylogenetic « relic ». *Evolution*, 1970, 24, 402-414.
[77] *Stern C.* Principles of human genetics. San Francisco-London, Freeman and Co., 1960.
[78] *Tobari J. N., Kojima K.* A study of spontaneous mutation rates at ten loci detected by starch gel electrophoresis in *Drosophila melanogaster*. *Genetics*, 1972, 70, 397-403.
[79] United Nations Scientific Committee on the Effects of Atomic Radiation. Sources and Effects of Ionizing Radiation. N.Y., United Nations, 1977.
[80] *Wright S.* Evolution in Mendelian populations. *Genetics*, 1931, 16, 97-159.
[81] *Wright S.* The genetical structure of populations. *Ann. Eugen.*, 1951, 15, 323-354.

SÉLECTION DIRIGÉE CONTRE LES VARIANTES ÉLECTROPHORÉTIQUES RARES DES PROTÉINES ET CADENCES DE LA MUTATION SPONTANÉE DANS LES POPULATIONS

Y. Altoukhov, V. Doukharev et L. Jivotovski

Une immense variabilité héréditaire des populations de différentes espèces a été découverte par l'électrophorèse des protéines en tant que marqueurs génétiques: jusqu'à 30 % et même plus des protéines étudiées, principalement de nature enzymatique, se sont trouvées polymorphes pour une hétérozygotie moyenne par locus de 5 à 15 % [20, 26].

Ces découvertes ont déclenché une discussion sur le sens biologique d'une telle variabilité et des mécanismes responsables de son maintien [1, 9, 22, 24, 25, 31, 32]. Selon un des points de vue, la variabilité observée suivant les loci protéiques peut s'expliquer par la neutralité sélective d'une importante partie des mutations *de novo* et par la dérive génétique. Selon d'autres, l'apport essentiel au maintien du polymorphisme biochimique est fourni par différentes formes de sélection équilibrée. La question des mutations *de novo* pour les gènes correspondants n'a pas fait l'objet d'une discussion ou, tout au moins, on considère qu'une partie d'entre elles a des avantages sélectifs.

Cependant, il faut reconnaître que ces modèles négligent de considérer un fait: dans les populations naturelles, on découvre toujours des protéines génétiquement monomorphes représentées dans les aires d'habitation des espèces par un type fortement prédominant [2, 8]. La gamme des fréquences des variantes aberrantes ($P \leqslant 0{,}05$-$0{,}01$) suggère qu'elles ne se maintiennent que grâce à l'équilibre entre la mutation spontanée et la sélection dirigée contre elles. Supposant que la sélection soit aussi dirigée contre les variantes rares de protéines polymorphes, nous pouvons subdiviser tout le pool des allèles dans les limites de l'espèce en deux groupes: les allèles ordinaires, anciens du point de vue évolutif, se rencontrant aux fréquences polymorphes et les allèles rares, jeunes du point de vue évolutif, non entrés dans le patrimoine génétique intégré et coadapté de l'espèce à cause de leur influence néfaste sur l'adaptabilité [1, 3].

Ces derniers temps, des preuves ont été obtenues en faveur de ce

© Издательство Наука, 1983.

modèle. L'étude des prises de sang et d'autres tissus de l'homme a montré que, dans le cas d'adaptabilité fortement diminuée (enfants prématurés, enfants porteurs de malformations graves et multiples), la fréquence des variantes électrophorétiques rares de protéines se trouve à peu près à un ordre de grandeur plus élevée que chez les nouveau-nés normaux et dans le système reproducteur de la population. Parfois, de telles variantes peuvent être interprétées comme les marqueurs de mutations *de novo*, c'est-à-dire apparues dans une génération donnée [3, 7, 12, 14]. Apparemment, ce résultat est fondamental pour la génétique des populations. Cependant, il faut reconnaître les difficultés auxquelles on se heurtera en étudiant les stades précoces de l'ontogénèse de l'homme, surtout en ce qui concerne le choix d'un matériel statistique fiable. Compte tenu de cette circonstance, nous avons décidé d'élargir le nombre d'espèces étudiées en prêtant une attention particulière à celles qui représentent le plus d'intérêt en tant que modèles.

Dans cet article, nous examinerons la distribution des variantes électrophorétiques rares de protéines dans les populations de pins sylvestres (*Pinus sylvestris* L.) aux différents stades ontogénétiques.

L'objet choisi présente un intérêt exceptionnel de deux points de vue. Premièrement, au cours des dernières années, on a décrit pour cette espèce et d'autres apparentées des systèmes isoenzymatiques variés décelés par électrophorèse, contrôlables génétiquement et se révélant de bonne heure dans l'ontogénèse [15, 16, 19, 27, 34, 35, 36, 37].

Deuxièmement, chose importante, le système reproducteur du pin, comme des autres gymnospermes, est tel qu'il permet d'étudier à la fois les phases diploïde et haploïde de développement. En utilisant les protéines comme marqueurs génétiques, on peut déterminer le génotype mère (aiguilles), le génotype des gamètes (endospermes), les génotypes des descendants (germes et pousses de différents âges). Il est évident que les variantes électrophorétiques rares découvertes dans les endospermes (macrogamétophytes) et absentes dans les aiguilles ne sont rien d'autre que le résultat de mutations (ou de recombinaisons) des gènes correspondants dans les gamètes maternels. De même, les variantes rares découvertes dans les germes et absentes dans les gamètes femelles sont le résultat de mutations apportées par le pollen.

Il est également évident que si, contre les variantes protéiques rares provenant de la mutation, agit la sélection, la fréquence de ces variantes doit être maximale justement au niveau des gamètes et diminuer dans la série endospermes → germes → pousses.

Ainsi, l'objet envisagé offre une possibilité unique non seulement d'évaleur les cadences de la mutation directement dans les gamètes, mais aussi de déterminer l'intensité de la sélection dirigée contre les mutations *de novo* aux différents stades de l'ontogénèse.

MATÉRIEL ET MÉTHODE

Quatre générations d'origine naturelle et éloignées des entreprises industrielles ont été étudiées.

La population Zvénigorodskaïa se trouve dans la région de Moscou, sur le territoire de la station biologique de l'Université de Moscou à Zvénigorod. La pinède à mousse verte est riche en airelles rouges. La futaie est à deux étages. L'âge moyen des arbres du premier étage est de 150 ans. Leur hauteur est de 30 m; leur diamètre, de 55 cm. Le second étage est représenté par des arbres dont l'âge moyen est de 45 ans; la hauteur, de 18 m; le diamètre, de 16 cm.

La population Mouchmar'inskaïa se trouve en R.S.S.A. de Mariis. L'âge moyen des arbres du premier étage est de 170 ans. Leur hauteur est de 30 m; leur diamètre, de 52 cm. Les arbres du deuxième étage ont les paramètres suivants: âge moyen 70 ans, hauteur 22 m, diamètre 24 cm. La pinède à mousse verte, riche en airelles rouges.

Les populations Pétrilovskaïa I (en sol sec) et Pétrilovskaïa II (sur tourbière) se trouvent à la limite des régions de Kalinine et de Pskov et sont directement voisines. La Pétrilovskaïa II est représentée par une pinède à ériophores et à sphaignes. L'âge moyen des arbres est de 50 ans. Leur hauteur est de 1 m; leur diamètre, de 4 cm. La Pétrilovskaïa I est représentée par une pinède à airelles rouges et noires. L'âge moyen des arbres est de 50 ans. Leur hauteur est de 19 m; leur diamètre, de 24 cm.

L'électrophorèse en bloc vertical sur gel de polyacrylamide suivie d'une coloration histochimique des endospermes, des germes et des aiguilles servait à identifier les enzymes suivantes: leucine-aminopeptidase (LAP; C. E. 3.4.11.1); glutamine-synthétase (GTS; C. E. 6.3.1.2); glucose-6-phosphate-déshydrogénase (G-6-DP; C. E. 1.1.1.49); 6-phosphogluconate-déshydrogénase (6-PGDH; C.E. 1.1.1.44); malate-déshydrogénase (MDH; C.E. 1.1.1.37); isocitrate-déshydrogénase (ICDH; C.E. 1.1.1.42); glutamo-oxalate-transaminase (GOT; C.E. 2.6.1.1.); superoxyde dismutase (SOD; C.E. 15.1.1.); 3-phosphoglycéraldéhyde-déshydrogénase (3-PGADH; C.E. 1.2.1.12); aldolase (Ald; C.E. 4.1.2.13); aconitase (Aco; C.E. 4.2.1.3.); alcool-déshydrogénase (ADH; C.E. 1.1.1.1); estérases non spécifiques (Est; C.E. 3.1.1.1.); phosphatase alcaline (PAl; C.E. 3.1.3.1); phosphatase acide (PAc; C.E. 3.1.3.2).

La malate-déshydrogénase, l'isocitrate-déshydrogénase, la 6-phosphogluconate-déshydrogénase, l'aldolase, l'aconitase, la 3-phosphoglycéraldéhyde-déshydrogénase, la glucose-6-phosphate-déshydrogénase, la superoxyde dismutase sont contrôlées chacune par un locus. Les phosphatases acide et alcaline, la leucine-aminopeptidase, la glutamo-oxalate-transaminase, l'alcool-déshydrogénase, la glutamine-synthétase sont gouvernées par deux gènes non alléliques. Les estérases non spécifiques sont contrôlées par quatre gènes non alléliques.

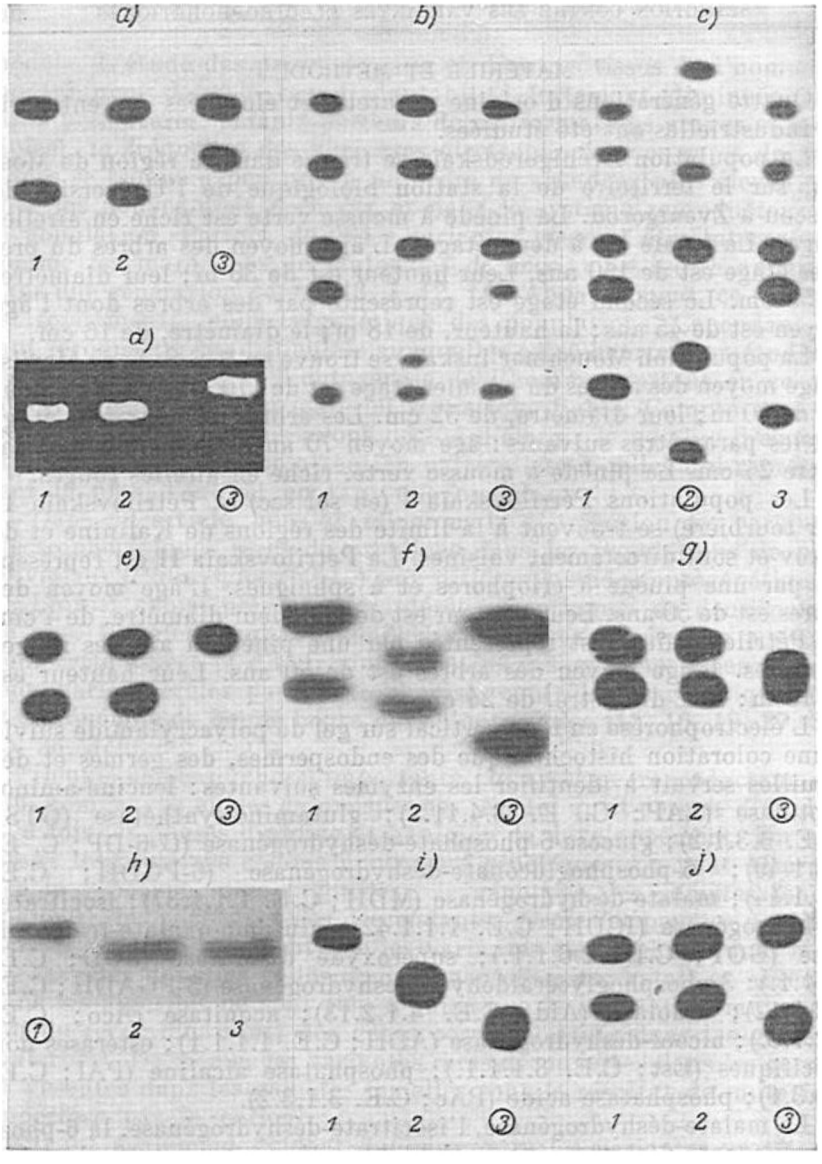


Fig. 1. Exemples de variantes protéiques électrophorétiques rares (chiffres encerclés) découvertes au cours de l'étude d'extraits aqueux d'endospermes (macrogamétophytes) du pin sylvestre:

a — phosphatase alcaline, mobilité modifiée; *b* — estérase C, allèle zéro; *c* — estérase A, mobilité modifiée; *d* — superoxyde dismutase, mobilité modifiée; *e* — phosphatase acide, allèle zéro; *f* — leucine-aminopeptidase, mobilité modifiée; *g* — glutamine-synthétase, mobilité modifiée; *h* — aconitase, mobilité modifiée; *i* — glucose-6-phosphate-déshydrogénase, mobilité modifiée; *j* — phosphatase acide, mobilité modifiée.

Sur tous les clichés, l'anode est en bas, la cathode en haut. Électrophorèse sur gel de polyacrylamide

Chaque graine était pesée, puis on ôtait sa pelure au moyen d'une aiguille spéciale. Les germes et les endospermes étaient utilisés pour l'analyse. L'extraction se faisait par la méthode de Feret [18]. Les méthodes d'identification des estérases non spécifiques, des phosphatases acide et alcaline, de la leucine-aminopeptidase ont été décrites par V. Doukharev [15, 16]. Toutes les autres enzymes ont été identifiées selon les techniques de Brewer [10]. Dans la population Pétrilovskaïa, les spectres électrophorétiques des enzymes étaient étudiés sur les endospermes et les germes prélevés sur 30 arbres. On prenait 100 graines sur chaque arbre. Dans les populations Zvénigorodskaïa et Mouchmar'inskaïa, on a étudié 100 germes et endospermes prélevés sur 25 arbres.

En outre, les graines recueillies des 25 arbres de la population Zvénigorodskaïa ont été semées et les spectres électrophorétiques des isoenzymes ont été étudiés sur les pousses pendant deux ans.

Les variantes électrophorétiques ont été classées en accord avec les principes de la génétique biochimique communément admis [1, 24, 25], compte tenu des différences individuelles tant dans la mobilité électrophorétique des produits protéiques correspondants que dans l'activité. Dans ce dernier cas, l'attention était accordée à la perte totale des zones correspondantes de l'activité enzymatique sur le zymogramme ou ce qu'on appelle allèles zéro (fig. 1).

Durant le travail à tous les stades étudiés de l'ontogénèse et pour tous les marqueurs ~700 000 loci-tests ont été effectués : plus de 430 000 identifications unitaires par électrophorèse.

PARTIE EXPÉRIMENTALE

Le plus grand nombre de loci identifiés par électrophorèse (23) ont été étudiés dans la population Pétrilovskaïa (*Tableau 1*). Cette analyse a été réalisée à deux stades ontogénétiques : les endospermes et les germes. Les populations Mouchmar'inskaïa et Zvénigorodskaïa se distinguent par un plus petit nombre de loci (10). Dans la population Zvénigorodskaïa on a réussi à suivre la distribution des variantes rares non seulement parmi les endospermes et les germes, mais aussi dans deux groupes de pousses, c'est-à-dire dans une période plus grande de l'ontogénèse (*Tableau 2*). Le trait commun à ces deux *Tableaux* est qu'on peut en tirer une information tant sur l'effectif des échantillons étudiés et sur la quantité de variantes électrophorétiques rares que sur le nombre d'allèles zéro et d'allèles issus de la mutation, ceux apportés par le pollen du nombre. Cette présentation du matériel permet

1° d'évaluer quantitativement la proportion entre la part de mutations conduisant à la perte de l'activité enzymatique de la molécule et la part de mutations modifiant sa mobilité électrophorétique ;

Variantes électrophorétiques rares identifiées

Loci isoenzymatiques étudiés	Populations étudiées, stades ontogénétiques							
	Pétrilovskaïa I							
	Endospermes				Germes			
		des variantes rares				des variantes rares		
	des loci-tests	avec perte d'activité	à mobilité modifiée	Total	des loci-tests	avec perte d'activité	à mobilité modifiée	Total
Polymorphes :								
Arylestérase	3 000	1	2	3	6 000	1	(1)	2
Carboxylestérase	3 000	1	3	4	6 000	1	1	2
Acétylestérase	3 000		2	2	6 000		2 (1)	2
Phosphatase alcaline A	3 000	1	1	2	6 000			
Phosphatase alcaline C	3 000		1	1	6 000		(1)	1
Phosphatase acide B	3 000		2	2	6 000			
Phosphatase acide C	3 000				6 000			
Leucine-aminopeptidase A	3 000	1	1	2	6 000	1		1
Leucine-aminopeptidase B	3 000		2	2	6 000		2 (1)	2
Superoxyde dismutase	3 000	1	1	2	6 000			
Glutamo-oxalate-transaminase B	3 000		1	1	6 000			
Malate-déshydrogénase	3 000		1	1	6 000			
Phosphogluconate-déshydrogénase	3 000		1	1	6 000			
Glucose-6-phosphate-déshydrogénase	3 000	1	1	2	6 000		1	1
Total pour les loci polymorphes	42 000	6	19	25	84 000	3	8	11
Monomorphes :								
Alcool-déshydrogénase A	3 000				6 000			
Alcool-déshydrogénase B	3 000		1	1	6 000		1	1
Glutamo-oxalate-transaminase A	3 000	1		1	6 000			
Glutamine-synthétase A	3 000				6 000			
Glutamine-synthétase B	3 000				6 000			
Aldolase	3 000		1	1	6 000			
Aconitase	3 000				6 000			
Isocitrate-déshydrogénase	3 000		1	1	6 000			
3-Phosphoglycéraldéhyde-déshydrogénase	3 000				6 000			
Total pour les loci monomorphes	27 000	1	3	4	54 000		1	1
Total pour les loci monomorphes et polymorphes	69 000	7	22	29	138 000	3	9	12

Remarque. Entre parenthèses, les variantes rares apportées par le pollen.

Tableau 1

dans les populations de pin sylveste Pétrilovskaïa

et types de variantes protéiques

Pétrilovskaïa II							
Endospermes				Germes			
des loci-tests	des variantes rares			des loci-tests	des variantes rares		
	avec perte d'activité	à mobilité modifiée	Total		avec perte d'activité	à mobilité modifiée	Total
3 000	3	3	6	6 000	2	1	3
3 000	2	3	5	6 000	1	2 (1)	3
3 000	2	4	6	6 000	1	2 (1)	3
3 000	1	2	3	6 000		1	1
3 000	1	1	2	6 000		(1)	1
3 000	1	1	2	6 000			
3 000	1	2	3	6 000	1		1
3 000	2	2	4	6 000	1	2 (1)	3
3 000	2	3	5	6 000	1	(1)	2
3 000	1	2	3	6 000		1	1
3 000		2	2	6 000			
3 000	2	1	3	6 000		1	1
3 000	1	1	2	6 000		(1)	1
3 000	1	2	3	6 000		2 (1)	2
42 000	20	29	49	84 000	7	15	22
3 000	1		1	6 000		(1)	1
3 000		1	1	6 000		1	1
3 000		1	1	6 000			
3 000		1	1	6 000			
3 000		1	1	6 000			
3 000				6 000			
3 000				6 000			
3 000				6 000			
3 000		1	1	6 000			
27 000	1	5	6	54 000		2	2
69 000	21	34	55	138 000	7	17	24

Variantes électrophorétiques rares identifiées dans les populations

Loci isoenzymatiques étudiés	Populations étudiées, stades ontogénétiques							
	Mouchmar'inskaïa							
	Endospermes				Germes			
	des loci-tests	de variantes rares			des loci-tests	de variantes rares		
		avec perte d'activité	à mobilité modifiée	Total		avec perte d'activité	à mobilité modifiée	Total
Arylestérase	2 500	1	2	3	5 000		1	1
Carboxylestérase	2 500	1	2	3	5 000	1	1	2
Acétylestérase	2 500	1	1	2	5 000		1	1
Carbcanhydrase	2 500	1	1	2	5 000		(1)	1
Phosphatase alcaline A	2 500		1	1	5 000		1	1
Phosphatase alcaline C	2 500		1	1	5 000		(1)	1
Phosphatase acide B	2 500		2	2	5 000			
Phosphatase acide C	2 500	1		1	5 000			
Leucine-aminopeptidase A	2 500		2	2	5 000		(1)	1
Leucine-aminopeptidase B	2 500	1	2	3	5 000	1		1
Total	25 000	6	14	20	50 000	2	7	9

Remarque. Entre parenthèses, les variantes apportées par le pollen. On n'a découvert

2° de faire des estimations différentes des cadences de la mutation pour ces groupes d'allèles, ainsi que de déterminer la vitesse moyenne de mutation par locus par gamète (et par individu);

3° de comparer la vitesse de mutation séparément pour les loci poly et monomorphes;

4° d'évaluer les coefficients de sélection dirigée contre les mutations *de novo* d'après les loci isoenzymatiques étudiés à des stades ontogénétiques allant des gamètes aux pousses de 12 mois;

5° de comparer la vitesse de mutation dans des populations différentes.

Cette analyse se fera plus facilement si l'on se réfère au *Tableau 3* qui fournit une généralisation des données des *Tableaux 1* et *2*. On voit que pour toutes les populations étudiées une même particularité est typique: la fréquence des variantes électrophorétiques rares parmi les endospermes est authentiquement plus élevée que parmi les

Tableau 2

de pins sylvestres Mouchmar'inskaïa et Zvénigorodskaïa

et types de variantes protéiques rares

Zvénigorodskaïa											
Endospermes				Germes				Pousses de 6 mois		Pousses de 1 an	
	de variantes rares				de variantes rares						
des loci-tests	avec perte d'activité	à mobilité modifiée	Total	des loci-tests	avec perte d'activité	à mobilité modifiée	Total	des loci-tests	des variantes rares	des loci-tests	des variantes rares
2 500		2	2	5 000		2 (1)	2	5 000	1	5 000	
2 500		3	3	5 000		(1)	1	5 000		5 000	
2 500	1	2	3	5 000	1		1	5 000		5 000	
2 500	1	2	3	5 000		1	1	5 000	1	5 000	
2 500		2	2	5 000	1	1	2	5 000		5 000	
2 500		1	1	5 000				5 000		5 000	
2 500		1	1	5 000		(1)	1	5 000		5 000	
2 500		2	2	5 000				5 000	1	5 000	
2 500	1	2	3	5 000	1	(1)	2	5 000	1	5 000	
2 500	2	1	3	5 000	1		1	5 000		5 000	
25 000	5	18	23	50 000	4	7	11	50 000	4	50 000	0

parmi les pousses que des variantes à mobilité électrophorétique modifiée.

germes. A ce stade de l'ontogénèse, les coefficients de sélection s'avèrent curieusement constants et élevés dans toutes les populations en accord avec ce qui est connu pour les létaux dominants ou, tout au moins, semiletaux. Les données sur la population Zvénigorodskaïa dont l'ontogénèse a été étudiée montrent exhaustivement qu'au stade de pousses de 12 mois on ne parvient déjà plus à découvrir des variantes rares et que, par conséquent, les coefficients de sélection pour toute la période envisagée de l'ontogénèse peuvent être évalués à 1. Ce résultat correspond parfaitement aux hypothèses formulées antérieurement et aux données expérimentales récentes discutées plus haut. Il fournit une preuve encore plus directe et convaincante de la forte pression de sélection sur les variantes électrophorétiques protéiques rares. Mais ceci signifie que de telles variantes se rapportent, pour la plupart, comme on le supposait, à la catégorie des mutations récentes dont seulement une très faible

Tableau 3

Fréquence ($1 \cdot 10^{-4}$) des variantes protéiques électrophorétiques rares survenant dans les gamètes maternels et coefficients de sélection dirigée contre elles dans les populations de pin sylvestre

Loci	Variantes identifiées	Populations étudiées								Total pour toutes les populations		
		Mouchmar'inskaïa		Zvénigorodskaïa		Pétrilovskaïa I		Pétrilovskaïa II				
		Endospermes	Germes	Endospermes	Germes	Endospermes	Germes	Endospermes	Germes	Endospermes	Germes	Coefficient de sélection
Polymorphes	Avec perte de l'activité	2,4	0,8	2,0	1,6	1,4	0,7	4,8	1,7	2,8	1,2	0,57
	A mobilité modifiée	5,6	1,6	7,2	1,2	4,5	0,9	6,9	1,9	6,0	1,4	0,77
	Total	8,0	2,4	9,2	2,8	5,9	1,7	11,7	3,6	8,8	2,6	0,70
Monomorphes	Avec perte de l'activité					0,4	0,0	0,4	0,0	0,4	0,0	1,00
	A mobilité modifiée					1,1	0,4	1,8	0,4	1,5	0,4	0,73
	Total					1,5	0,4	2,2	0,4	1,9	0,4	0,79
Total pour les deux groupes de loci pour les variantes avec perte de l'activité						1,0	0,4	3,0	1,0	2,1	0,8	0,62
Total pour les deux groupes de loci pour les variantes à mobilité modifiée						3,2	0,7	4,9	2,2	4,7	1,1	0,77
Total pour tous les loci et variantes						4,2	1,1	8,0	3,2	7,7	2,0	0,74

fraction pénètrent dans le système reproducteur de la population et existent dans un petit nombre de générations comme étant plus ou moins sélectivement neutres *.

On comprend qu'en présence d'une si forte pression de sélection, toutes les évaluations indirectes des cadences de la mutation spontanée faites jusqu'à présent d'après les données sur les variantes protéiques rares découvertes aux stades ontogénétiques avancés, doivent être rectifiées. On y avait déjà pensé auparavant et les résultats du présent travail attestent la même chose: l'estimation globale pour tous les loci et populations de la vitesse de mutation lors de l'étude des gamètes est de $7{,}7 \cdot 10^{-4}$, alors qu'à l'analyse des germes elle est de $2{,}0 \cdot 10^{-4}$. Si l'on considère la population Zvénigorodskaïa (v. *Tableau 2*), l'électrophorèse des protéines des aiguilles des pousses met en évidence une fréquence des variantes rares encore plus faible: $8{,}0 \cdot 10^{-5}$.

Les données envisagées montrent également que la fréquence des variantes protéiques rares diffère significativement pour les loci poly et monomorphes.

Ce fait admet deux explications ne s'excluant pas l'une l'autre.

1° Une cadence de la mutation (ou de la recombinaison) est plus élevée pour les loci polymorphes.

2° Une sélection dirigée contre les mutations *de novo* et les recombinaisons entre cistrons est plus forte pour les loci monomorphes.

Si ces suppositions sont justes, on peut attendre une corrélation positive entre l'hétérozygotie du locus (ou le nombre d'allèles « fréquents » qu'il comprend) et le nombre d'allèles rares découverts par électrophorèse. En effet, une telle corrélation est facile à établir, ce que l'on montre ici sur l'exemple de la population Pétrilovskaïa caractérisée d'après le plus grand nombre de loci (fig. 2).

Etant donné que la probabilité d'une recombinaison entre cistrons dans les loci monomorphes (possible seulement en cas de crossing-over inégal) doit être de beaucoup inférieure à celle des loci polymorphes, la fréquence gamétique des variantes rares pour le premier groupe de gènes sera la plus proche de l'estimation de la vitesse de mutation vraie ($1{,}9 \cdot 10^{-4}$, v. *Tableau 3*) et la fréquence correspondante pour le second groupe de gènes ($8{,}8 \cdot 10^{-4}$, v. *Tableau 3*) reflète aussi bien la pression de mutation propre que la recombinaison intragénique. En admettant l'identité des cadences de mutation des gènes poly et monomorphes, on peut évaluer la fréquence des recombinaisons. Comme on voit, la valeur obtenue ($8{,}8 - 1{,}9 = 6{,}9 \cdot 10^{-4}$, v. *Tableau 3*) ne diffère pas significativement des évaluations faites auparavant [11, 17, 23, 33, 38, 39], surtout si

* Ces derniers temps, nous avons obtenu des données permettant d'affirmer que chez l'homme même les variantes électrophorétiques rares héritées des protéines telles que les estérases érythrocytaires non spécifiques sont responsables de différents états pathologiques.

l'on tient compte de ce que la pression de sélection dirigée contre les allèles recombinants n'était pas prise en considération.

En ce qui concerne l'influence des variantes nouvellement survenues sur la viabilité, il serait difficile de s'attendre ici à des dif-

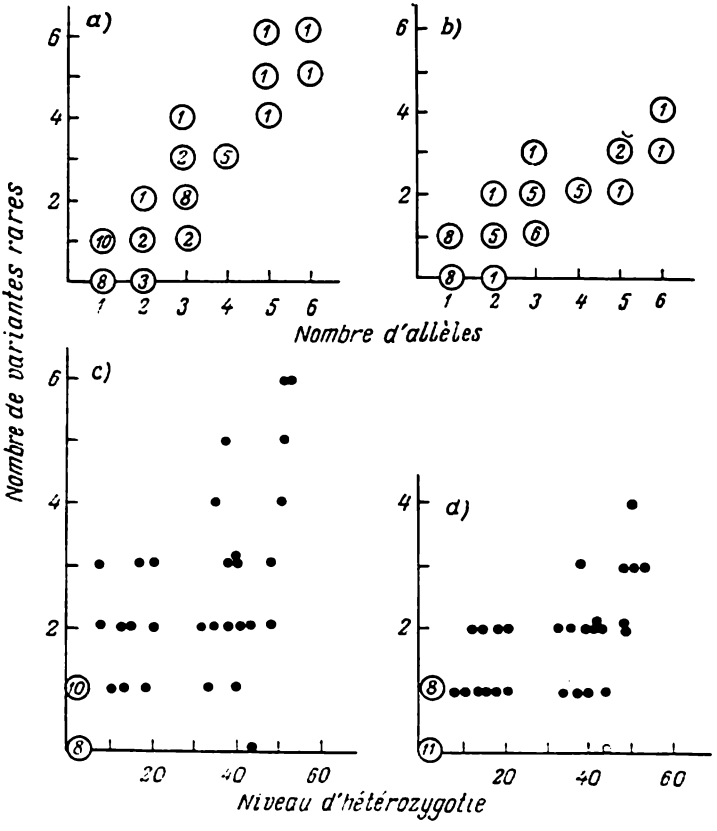


Fig. 2. Corrélation entre le nombre de variantes protéiques électrophorétiques nouvellement formées, le nombre d'allèles dans le locus (*a* et *b*) et le niveau d'hétérozygotie (*c* et *d*):

b et *d* — seulement pour les variantes à mobilité électrophorétique modifiée: $r_b = 0{,}91$ ($P < 0{,}001$); $r_d = 0{,}74$ ($P < 0{,}001$); *a* et *c* pour toutes les variantes identifiées, y compris les allèles zéro: $r_a = 0{,}84$ ($P < 0{,}001$); $r_c = 0{,}72$ ($P < 0{,}001$)

férences significatives entre les allèles de mutation et ceux de recombinaison (mutation-like events), dans les deux cas l'effet doit être délétère. Outre les différences significatives en rapport avec une plus grande fréquence des variantes rares aux stades ontogénétiques précoces, une autre preuve de l'influence néfaste sur la viabilité de toutes les mutations nouvelles identifiées par l'électrophorèse

est fournie par les résultats de l'étude de la distribution du poids des endospermes, si l'on considère ce caractère comme l'indice d'un développement normal [33, 34]. Des données de ce genre (fig. 3) témoignent de différences significatives dans le poids des endospermes, porteurs de variantes électrophorétiques rares, et des endospermes ayant un génotype normal. Comme on pouvait s'y attendre en accord avec le modèle bien fondé [6, 13], les endospermes, porteurs

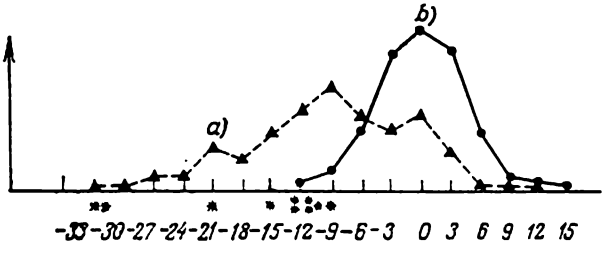


Fig. 3. Distribution de la masse des endospermes (macrogamétophytes): *a* — variabilité de la masse des endospermes portant des variantes protéiques électrophorétiques nouvellement formées ($\bar{X} = -9{,}1$; $\sigma = 8{,}89$); *b* — variabilité de la masse des endospermes normaux (non mutants) ($\bar{X} = 0$; $\sigma = 4{,}31$).
Les valeurs de la masse des endospermes portant des variantes protéiques mutantes des loci monomorphes sont marquées par des croix. Les distributions sont construites pour les écarts individuels par la masse de chaque endosperme des moyennes correspondantes typiques des arbres mères

de mutations *de novo*, ont une masse authentiquement plus petite que les endospermes du groupe témoin. Nous avons découvert le même tableau lors d'une étude analogue des germes et des pousses.

Ce fait est d'une grande importance pratique: quand on évalue l'intensité de la mutation dans les populations des pins ou d'autres gymnospermes, surtout à cause des problèmes de la pollution de l'environnement, il n'est pas nécessaire de procéder à un screening de masse, il suffit de se borner à l'électrophorèse des endospermes (graines) s'écartant fortement par leur masse de la normale et de définir leur part dans la population.

Les données de la figure 3 permettent aussi de caractériser la distribution par leur masse des endospermes portant les mutations des gènes monomorphes. On voit que les valeurs correspondantes s'écartent du centre de distribution vers la gauche. Ce résultat témoigne en faveur de la supposition qu'il existe une pression de sélection plus forte contre les mutations affectant la structure des protéines monomorphes. Apparemment, pour trancher cette question, il faut encore rassembler bien des données expérimentales.

Le présent article a encore un autre résultat: la différence évidente dans les cadences de la mutation entre deux populations Pétri-

lovskaïa liées par une origine commune. La vitesse de mutation s'avère sensiblement plus élevée dans la population Pétrilovskaïa II habitant dans des conditions écologiques exceptionnellement dures.

DISCUSSION

L'article présent qui complète les résultats publiés auparavant, donne encore un exemple de l'efficacité de l'approche proposée du screening électrophorétique des mutations *de novo* modifiant la structure de la protéine. Cependant, à la différence des recherches précédentes effectuées sur un matériel réduit et limitées aux stades précoces de l'ontogénèse postnatale et aux stades assez tardifs du développement prénatal, cet ouvrage renferme beaucoup de matériel expérimental et embrasse une période plus grande de l'ontogénèse, y compris le stade des gamètes. Cette dernière circonstance nous semble particulièrement importante, car elle fournit des preuves directes et indiscutables de ce que les variantes électrophorétiques rares identifiées sont issues de la mutation et que la sélection agit contre elles.

Ces résultats nouveaux sont d'une importance fondamentale pour la génétique des populations et évolutive.

Premièrement, l'inacceptabilité de la thèse d'une neutralité sélective (ou d'une faible nocivité) de la plupart des mutations *de novo* pour les gènes contrôlant la structure des protéines. Nos données attestent que les coefficients de sélection dirigée contre de telles mutations, à une étape de développement s'étendant depuis le stade des gamètes jusqu'à celui du germe, dépassent en moyenne 70 % pour toutes les populations et tous les loci étudiés et atteignent 100 % chez les pousses de 12 mois, comme nous l'avons vu dans le cas de la population Zvénigorodskaïa.

Deuxièmement, il devient évident que toutes les évaluations des cadences de la mutation spontanée faites jusqu'alors aux stades tardifs de l'ontogénèse et basées sur la conjecture d'une neutralité sélective des variantes protéiques rares, sont fausses. Aussi la conclusion déduite précédemment que l'évaluation de la cadence de la mutation pour les gènes correspondants doit être d'un ordre de grandeur plus élevée [4], trouve-t-elle sa confirmation dans les faits ici présentés. Cependant, on ne sait pas encore dans quelle mesure une telle cadence de la mutation peut être considérée comme historiquement constituée.

Les matériaux expérimentaux examinés soulèvent d'autres questions importantes. Comment est distribuée la pression de sélection contre les mutations *de novo* entre les différents loci et les stades de l'ontogénèse dans l'intervalle du stade des gamètes jusqu'au stade des germes avancés? Comment expliquer du point de vue biochimique l'influence délétère des mutations identifiées par électrophorèse sur l'embryogénèse et surtout sur le stade des gamètes?

En ce qui concerne le groupe d'allèles zéro, c'est-à-dire des mutations inhibant la synthèse du polypeptide ou son activité enzymatique, dans toute une série de cas leur influence néfaste sur l'adaptabilité a été démontrée en étudiant les troubles du métabolisme en rapport avec la perte de l'activité des enzymes, ainsi que des protéines de transport ou réceptrices chez l'homme [29, 30].

A ce propos, il convient d'évaluer le rapport des mutations « nulles » et « électrophorétiques » survenant spontanément dans les cellules germinales. Selon Mukai et Cockerham [28], ce rapport est de 5 : 1 chez la drosophile, c'est-à-dire que les allèles zéro correspondant formellement à la classe de mutations récessives, sont 5 fois plus fréquents que les mutations responsables d'une variation de la mobilité électrophorétique. Cependant, chez le pin, le tableau découvert est différent : la part des allèles zéro dans le groupe de loci biochimiques étudié par nous ne constitue en moyenne que le tiers du nombre total des mutations *de novo*, alors que les autres deux tiers se rattachent au groupe de mutations faisant varier la mobilité électrophorétique de la molécule protéique.

Le mécanisme de l'action des mutations « électrophorétiques » délétères exige une étude spéciale, mais le fait même ne soulève pas de doutes. Bien plus, dans le cas des loci polymorphes les coefficients de sélection dirigée contre ces mutations sont même plus importants que ceux observés pour le cas des allèles zéro (v. *Tableau 3*).

Et, enfin, il reste une chose à discuter. Elle se rattache à la nécessité de répondre à la question : dans quelle mesure les régulations ontogénétiques possibles de l'expression des loci étudiés peuvent influencer nos évaluations ? Nous estimons que des doutes de ce genre, même s'ils sont possibles, ne concernent que les variantes à déficit complet d'une activité enzymatique identifiées lors de l'étude du stade haploïde. En ce qui concerne les nouvelles variantes à mobilité électrophorétique modifiée, l'identification simultanée de tels génotypes dans l'endosperme et le germe dissipe tous les doutes sur leur véritable nature.

RÉFÉRENCES

[1] *Алтухов Ю. П.* Популяционная генетика рыб. М., Пищевая промышленность, 1974, 245 с.

[2] *Алтухов Ю. П.* О соотношении моно- и полиморфизма гемоглобинов в микроэволюции рыб. *Докл. АН СССР*, 1969, 189, 5, 1115.

[3] *Altukhov Yu. P.* Environmental conditions and genetic monitoring of populations. In: Well-being of mankind and genetics. Proc. XIV Internat. Congr. Genet., 1. Moscow, MIR Publ., 1980, 238.

[4] *Altukhov Yu. P.* Mutation rate in man under selection pressure against rare protein variants. In: Abstr. 11th Annual Meeting of the European environmental mutagen society. July 6-11, 1981. Budapest, 1981, 11.

[5] *Алтухов Ю. П., Ботвиньев О. К., Курбатова О. Л.* Популяционно-генетический подход к проблеме неспецифической биологической устойчивости человеческого организма. Сообщение 1. Постановка проблемы

и обоснование подхода. Параметры распределений антропометрических признаков новорожденных и грудных детей в норме и патологии. *Генетика*, 1979, 15, 2, 352.

[6] *Altukhov Yu. P., Dubinin N. P., Khilchevskaya R. I.* Genetic monitoring of human populations. In: Abstr. V International Congress Human Genetics. Mexico, 1976, 176.

[7] *Алтухов Ю. П., Дуброва Ю. Е.* Биохимический полиморфизм популяций и его биологическое значение. *Успехи соврем. биол.*, 1981, 91, 3, 467.

[8] *Алтухов Ю. П., Салменкова Е. А., Омельченко В. Т., Сачко Г. Д., Слынько В. И.* О числе мономорфных и полиморфных локусов в популяции кеты, *Oncorhynchus keta* Walb. — одного из тетраплоидных видов тихоокеанских лососей. *Генетика*, 1972, 8, 2, 251.

[9] *Ayala F. J.* Molecular evolution. Sunderland (Mass.), Sinauer Ass. Publ., 1976, 277.

[10] *Brewer G. J.* An introduction to isozyme techniques. N.Y.-L., Acad. Press, 1970, 195.

[11] *Chovnick A., Ballantyne G. H., Baillie D. L., Holm D. C.* Gene conversion in higher organisms: half-tetrad analysis of recombination with the rosy cistron of *Drosophila melanogaster*. *Genetics*, 1970, 66, 315.

[12] *Dubinin N. P., Altukhov Yu. P.* Gene mutations (*de novo*) found in electrophoretic studies of blood proteins of infants with anomalous development. *Proc. Nat. Acad. Sci. U.S.A.*, 1979, 76, 5226.

[13] *Дубинин Н. П., Алтухов Ю. П.* Мониторинг генетических последствий загрязнения окружающей среды: общий подход. В кн.: Генетические последствия загрязнений окружающей среды. М., Мысль, 1977, 14-45.

[14] *Дубинин Н. П., Алтухов Ю. П., Сусков И. И., Хильчевская Р. И., Афанасьев К. И., Малинина Т. В., Шурхал А. В., Большакова И. К., Курбатова О. Л., Прохоровская В. Д.* Экспериментальное обоснование принципов мониторинга генных мутаций. *Докл. АН СССР*, 1978, 243, 5, 1313.

[15] *Духарев В. А.* Состав эстераз у сосны обыкновенной в разновозрастных древостоях. *Лесоведение*, 1978, 6, 45.

[16] *Духарев В. А.* Биохимический полиморфизм в популяциях сосны обыкновенной (*Pinus sylvestris* L.) Автореф. дис. М., Ин-т общей генет. АН СССР, 1979, 16.

[17] *Eanes W. F., Koehn R. K.* The correlation of rare alleles with heterozygosity: determination of the correlation for the neutral models. *Genet. Res.*, 1977, 29, 223.

[18] *Feret P. P., Bergmann F.* Gel electrophoresis of proteins and enzymes. In: Modern methods in forest genetics. N.Y., Springer-Verlag, 1976, 49.

[19] *Guries R. P., Ledig T.* Inheritance of some polymorphic isoenzymes in pitch pine (*Pinus ridida* Mill.). *J. Heredity*, 1978, 40, 27.

[20] *Harris H.* Enzyme polymorphisms in man. *Proc. Roy. Soc. Lond.*, 1966, Ser. B164, 298.

[21] *Karn M. N., Penrose L. S.* Birth weight and gestation time in relation to maternal age, parity and infant survival. *Ann. Eugen.*, 1951, 47, 147.

[22] *Kimura M.* The neutral theory of molecular evolution. *Sci. Amer.*, 1979, 241, 94.

[23] *Koehn R. K., Eanes W. F.* An analysis of allelic diversity in natural populations of *Drosophila*: the correlation of rare alleles with heterozygosity. In: Population genetics and ecology. N.Y., Acad. Press, 1976, 377.

[24] *Корочкин Л. И., Серов О. Л., Пудовкин А. И., Ароштам А. А., Боркин Л. Я., Малецкий С. И., Полякова Е. В., Манченко Г. П.* Генетика изоферментов. М., Наука, 1977, 278 с.

[25] *Lewontin R. C.* The genetic basis of evolutionary change. N.Y.-London, Columbia Univ. Press, 1974.

[26] *Lewontin R. C., Hubby J. L.* A molecular approach the study of genic

heterosozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics*, 1966, 54, 595.

[27] *Morris R. W.*, *Critchfield W. B.*, *Fowler D. P.* The putative austrian × red pine hybrid: a test of paternity based on allelic variation at enzyme-specifying loci. *Silvae Genet.*, 1980, 29, 93.

[28] *Mukai T.*, *Cockerham C. C.* Spontaneous mutation rates at enzyme loci in *Drosophila melanogaster*. *Proc. Nat. Acad. Sci. U.S.A.*, 1977, 74, 2514.

[29] *Neel J. V.* Some considerations pertinent to monitoring human populations for changing mutation rates. In: Well-being of mankind and genetics. Proc. XIV Internat. Congr. Genet., 1, Moskow, Mir Publ., 1980, 225-238.

[30] *Neel J. V.* Mutation and disease in man. *Canad. J. Genet. a. Cytol.*, 1978, 20, 295.

[31] *Nei M.* Molecular population genetics and evolution. Amsterdam, North-Holland Publ. Co., 1975, 278 p.

[32] *Nevo E.* Genetic variation in natural populations: patterns and theory. *Theoret. Popul. Biol.*, 1978, 13, 121.

[33] *Ohno S.*, *Stenius C.*, *Christian L.*, *Schipmann G. De novo* mutation-like events observed at the 6PGD locus of the Japanese quail and the principle of polymorphism breeding more polymorphism. *Biochem. Genet.*, 1969, 3, 1969.

[34] *O'Malley D. M.*, *Allendorf F. W.*, *Blake G. M.* Inheritance of isozyme variation and heterozygosity in *Pinus ponderosa*. *Biochem. Genet.*, 1979, 17, 233.

[35] *Rudin D.* Inheritance of glutamate-oxalate-transaminases (GOT) from needles and endosperms of *Pinus sylvestris*. *Hereditas*, 1975, 80, 296.

[36] *Rudin D.* Leucino-amino-peptidase (LAP) from needles and macrogametophytes of *Pinus sylvestris*. L. Inheritance of allozymes. *Hereditas*, 1977, 85, 219.

[37] *Rudin D.*, *Ekberg I.* Linkage studies in *Pinus sylvestris* L. using macrogametophyte allozymes. *Silvae Genet.*, 1978, 27, 1, 1.

[38] *Tsuno K.* Studies on mutation at esterase loci in *Drosophila virilis*. I. Spontaneous mutation rates and newly arisen variants. *Japan. J. Genet.*, 1981, 56, 155.

[39] *Watt W. B.* Intragenic recombination as a source of population genetic variability. *Amer. Natur.*, 1972, 106, 737.

[40] *Wright J. E.*, *Atherton L.* Genetic control of interallelic recombination at the LDH B locus in brook trout. *Genetics*, 1968, 60, 240.

APPROCHE DU PROBLÈME DE LA RÉSISTANCE BIOLOGIQUE NON SPÉCIFIQUE DE L'ORGANISME HUMAIN AU NIVEAU DE LA GÉNÉTIQUE DES POPULATIONS

Communication I. Position du problème et justification de l'approche. Paramètres des distributions des caractères anthropométriques de nouveau-nés et de nourrissons sains et malades

Y. Altoukhov, O. Botviniev et O. Kourbatova

Le problème de la résistance de l'organisme humain aux maladies est ordinairement étudié sans tenir compte de ce que chaque individu fait partie d'une population possédant un lot d'adaptations génétiques au milieu concret dans lequel s'est déroulée son histoire naturelle. Tout au moins certaines de ces adaptations sont contrôlées par des gènes mendéliens et le cas de l'anémie drépanocytaire en est l'exemple le plus éclatant. Une telle résistance peut être appelée *spécifique*.

A l'opposé, la *résistance non spécifique* doit être rattachée aux influences conjuguées de nombreux gènes dont chacun par lui-même peut être presque neutre, mais associé à d'autres gènes, il est capable d'influencer la variabilité.

Si un tel système adaptatif existe chez l'homme, historiquement ce fait doit impliquer inévitablement des corrélations entre la structure intégrale du génotype, les particularités constitutionnelles de l'organisme et ses possibilités homéostatiques de résister à des maladies diverses et à d'autres perturbations du milieu tant intérieur qu'extérieur.

En effet, nous avions déjà [1] découvert cette dépendance. Il s'est avéré que

1° les enfants nouveau-nés, s'écartant au minimum de la moyenne de la population par la totalité des caractères anthropométriques distribués normalement, ont beaucoup plus rarement différents troubles morbides que les enfants s'écartant du centre de distribution de ces mêmes caractères vers la droite et surtout vers la gauche;

2° derrière les phénotypes morphologiquement « moyens », il y a des associations de génotypes se rencontrant aux fréquences intermédiaires, alors que les génotypes « fréquents » et, ce qui est particulièrement important, les génotypes « rares » sont prédominants parmi les individus les plus tarés s'écartant morphologiquement de la normale;

© Издательство Наука, 1979.

3° à cette distribution des génotypes correspond une structure déterminée des mariages reflétant les modifications survenues dans les orientations historiquement constituées et dans l'intensité des migrations.

Ainsi, nous avons mis en évidence chez l'être humain des associations génotypiques coadaptées cachant différents types morphologiques (« norme adaptative ») et qui peuvent être détruites par croisement consanguin (inbreeding) ou par accouplement non consanguin (outbreeding). La présente communication ouvre une série de publications qui déploient et justifient cette approche plus en détail. Dans cet article, nous examinerons les particularités des distributions statistiques de deux caractères quantitatifs, le poids et la taille à la naissance, dans plusieurs groupes d'enfants malades et d'enfants sains.

MATÉRIEL ET MÉTHODE

Dans deux maternités et deux hôpitaux cliniques de Moscou, nous avons recueilli des informations sur des enfants et leurs parents d'après un programme spécialement élaboré et incluant les points suivants: données cliniques et anamnestiques (somatométrie et somatoscopie compte tenu des « grandes » et « petites » malformations, caractères anthropométriques: poids et taille à la naissance); groupes sanguins ABO et Rhésus; données généalogiques et démographiques. Ont été utilisés aussi bien les résultats de l'observation clinique immédiate que les données des archives sur plusieurs années.

Tous les enfants ont été divisés en 5 groupes:

1° les nouveau-nés pratiquement sains observés en maternité (témoins): 396 sujets;

2° les nouveau-nés malades observés en maternité: 192 sujets;

3° les nouveau-nés malades observés en hôpital: 568 sujets;

4° les nourrissons malades observés en hôpital: 381 sujets;

5° les nouveau-nés et les nourrissons décédés en hôpital: 265 sujets.

La structure de la morbidité et les autres particularités typiques de chaque groupe d'enfants et de leurs parents seront décrites dans les communications correspondantes.

Pour le dépouillement et l'interprétation des données ont été utilisées les méthodes de la statistique variationnelle, des analyses corrélative et dispersionnelle, les procédés et méthodes de la génétique des populations.

RÉSULTATS

Les figures 1 et 2 montrent les distributions de 2 caractères quantitatifs, le poids et la taille à la naissance, dans les 5 groupes d'enfants comparés. Les principaux paramètres statistiques de ces distributions (moyenne, écart type, indices d'asymétrie et d'excès) sont

présentés au *Tableau 1*. On voit clairement que les distributions comparées diffèrent pour tous les paramètres estimés:

1° les nouveau-nés sains se caractérisent, quant à la taille et au poids, par une moyenne maximale, une variance minimale et

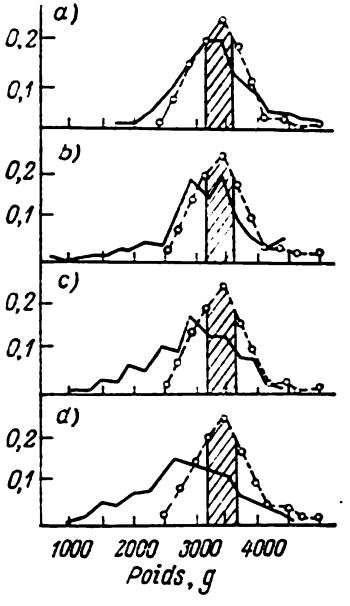


Fig. 1. Distribution du poids à la naissance dans un groupe de nouveau-nés malades (*a*) et de nourrissons malades (*b*) observés en hôpital, de nouveau-nés malades observés en maternité (*c*) et dans un groupe d'enfants décédés en hôpital (*d*)

Le pointillé indique la distribution dans le groupe témoin de nouveau-nés sains. Le domaine de « norme adaptative » est hachuré: $M \pm 1/2\,\sigma$

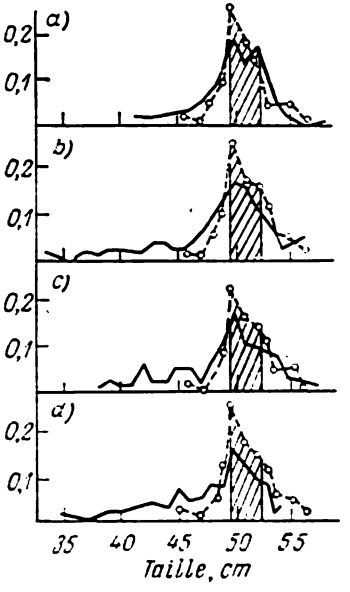


Fig. 2. Distribution de la taille à la naissance parmi les enfants malades et sains. Mêmes désignations que sur la figure 1

des indices d'asymétrie et d'excès ne se distinguant pas de façon certaine des indices correspondants pour une distribution normale;

2° les enfants décédés se caractérisent, pour les deux caractères, par une moyenne minimale, par une variance maximale, par une asymétrie prononcée à droite, particulièrement nette pour la distribution de la taille;

3° d'après les caractéristiques comparées, les autres groupes occupent une place intermédiaire entre les enfants sains et ceux qui sont morts, et ils se disposent par ordre d'altération des caractéristiques de la manière suivante: nouveau-nés malades à l'hôpital, nourrissons malades, nouveau-nés malades en maternité.

Tableau 1

Principaux paramètres statistiques des distributions du poids et de la taille des enfants sains et malades

Groupe d'enfants	N	Poids à la naissance, g						
		M	m	σ	γ_A	m_A	γ_E	m_E
Nouveau-nés sains	396	3458	21	428	0,3125	0,1231	3,2185	0,2462
Nouveau-nés malades observés en hôpital	556	3365	23	520	0,1250	0,1101	3,0412	0,1871
Nourrissons malades	369	3181	40	658	—0,5134 *	0,1480	3,6930	0,2980
Nouveau-nés malades observés en maternité	192	3084	51	707	—0,4004	0,1772	2,6882	0,3545
Morts	266	2834	46	754	—0,1477	0,1502	2,7631	0,3004

Taille à la naissance, cm							r	η^2	F_ξ
M	m	σ	γ_A	m_A	γ_E	m_E			
51,16	0,09	1,89	0,2812	0,1228	2,8829	0,2456	0,8102 ***	0,6742 ***	2,30 *
50,78	0,11	2,45	—0,5517 *	0,1105	4,3311 *	0,2211	0,8612 ***	0,7741 ***	6,50 ***
49,88	0,23	3,73	—1,6299 *	0,1516	8,9433 *	0,3032	0,8624 ***	0,7609 ***	1,81 *
49,30	0,28	3,87	—0,6831 *	0,1768	2,9253	0,3536	0,9178 ***	0,8539 ***	1,08
48,00	0,28	4,27	—0,8968 *	0,1629	3,4856	0,3259	0,8924 ***	0,8160 ***	1,85 *

Remarque. N est le nombre d'enfants étudiés ;
M la moyenne, m son erreur ;
σ l'écart type ;
γ_A l'indice d'asymétrie, m_A son erreur ;
γ_E l'indice d'excès, m_E son erreur ;
r le coefficient de corrélation ;
η^2 l'indice de liaison curviligne ;
F_ξ le critère de curvilinéarité.
Critère de certitude au niveau
* $P > 0,95$ et *** $P > 0,999$.

Il est donc évident que le domaine de « norme adaptative » doit être choisi non pas d'après la moyenne pour toute la population de nouveau-nés, mais à proximité de la moyenne pour les nouveau-nés sains. Prenant par convention $^1/_2\sigma$ à droite et à gauche du centre

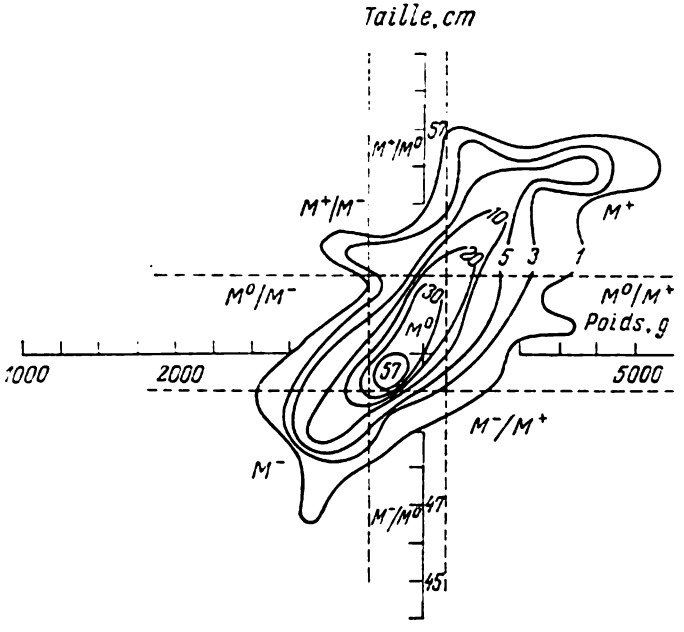


Fig. 3. Distribution conjointe du poids et de la taille à la naissance dans un groupe d'enfants sains et principe de classification des phénotypes suivant deux caractères anthropométriques (v. le texte):

Les chiffres sur les isolignes figurent le nombre d'enfants entrant dans la zone de grille de corrélation, par le milieu de laquelle passe l'isoligne

de la distribution correspondante, nous obtenons le domaine de « norme adaptative » avec les paramètres suivants: d'après le poids, de 3250 à 3650 g; d'après la taille, de 50 à 52 cm.

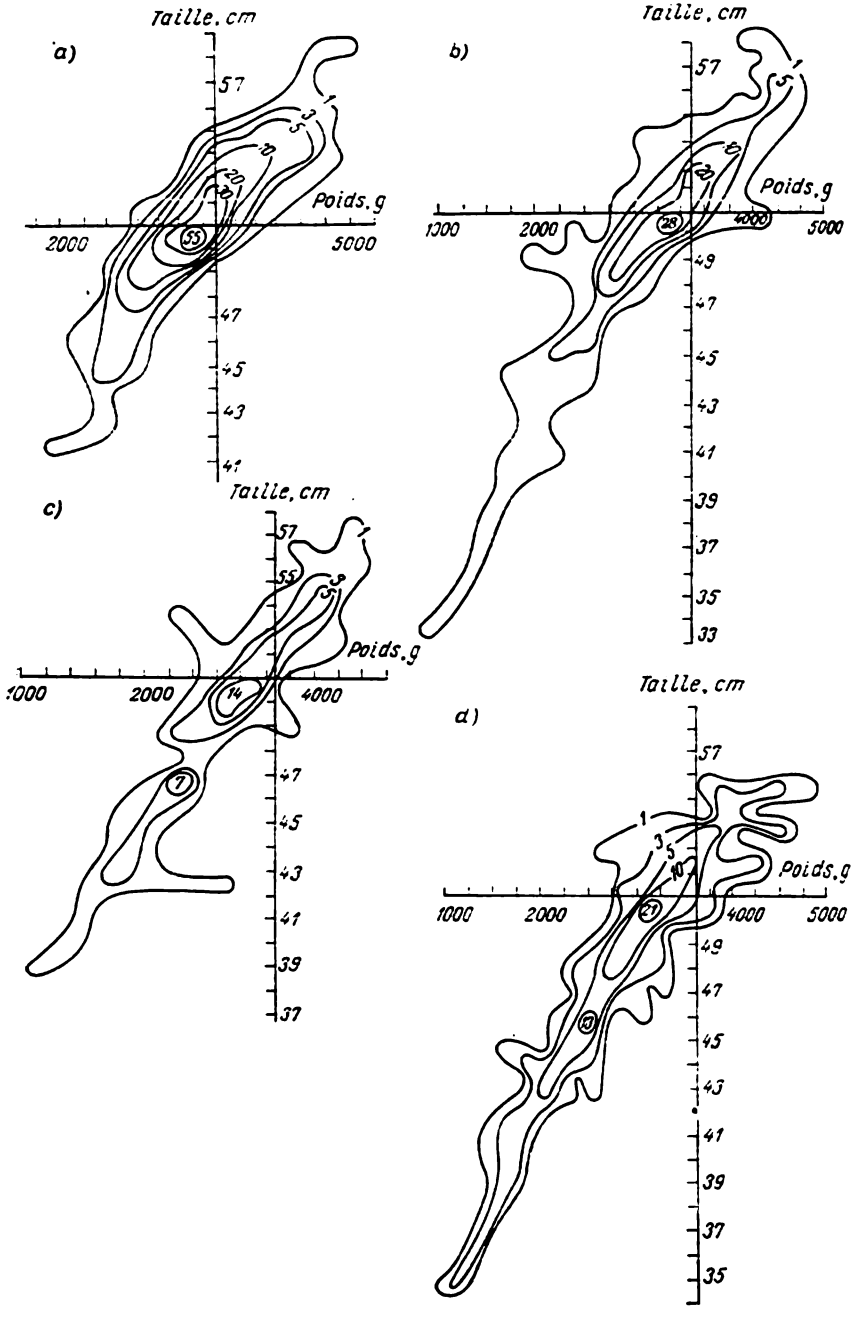
Ce domaine est hachuré sur les figures 1 et 2 et désigné M^0 pour chaque caractère pris en particulier. Le nombre d'enfants qu'il embrasse, diminue dans l'ordre suivant: enfants sains, nouveau-nés malades en hôpital, nourrissons malades, nouveau-nés malades en maternité, enfants décédés. La redistribution du nombre des enfants dans ces groupes a lieu principalement aux dépens de la diminution du domaine central (M^0) et du domaine droit extrême (M^+) et du fait de fréquences cumulées dans le domaine gauche (M^-).

Cependant, quand on examine deux caractères et plus, le principe de la séparation du domaine « de norme adaptative » sera plus

Tableau 2

Distribution des enfants sains et des enfants malades d'après les zones des champs de corrélation

Groupe d'enfants		Zone du champ de corrélation									
		Total	M^0	$M+$	$M-$	M^0/M^-	$M^0/M+$	M^-/M^0	$M+/M^0$	$M+/M^-$	$M^-/M+$
Nouveau-nés sains	Nombre	396	145	85	55	57	38	6	5	4	1
	%	100	36,6	21,5	13,9	14,4	9,6	1,5	1,3	1,0	0,2
	♂/♀	1,22	1,42	1,66	0,57	0,96	1,11	1,00	♂♂	♂♀	♀
Nouveau-nés malades observés en hôpital	Nombre	568	158	101	129	108	51	4	16	1	
	%	100	27,8	17,8	22,7	19,0	9,0	0,7	2,8	0,2	
	♂/♀	1,18	1,19	1,66	1,15	1,08	0,89	1,00	1,00	♀	
Nourrissons malades	Nombre	381	84	58	152	53	19	2	11	2	
	%	100	22,0	15,2	39,9	13,9	5,0	0,5	2,9	0,5	
	♂/♀	1,23	1,54	2,05	0,77	1,79	1,11	♀♀	4,5	1,0	
Nouveau-nés malades observés en maternité	Nombre	192	36	37	78	28	7	2	3	1	
	%	100	18,7	19,3	40,6	14,6	3,6	1,0	1,6	0,5	
	♂/♀	1,28	1,57	2,70	0,73	1,00	♂♂	♀♂	2,00	♂	
Morts	Nombre	265	34	27	152	34	10	2	4	2	
	%	100	12,8	10,2	57,4	12,8	3,8	0,7	1,5	0,7	
	♂/♀	1,39	1,27	2,37	1,34	1,00	1,50	1,00	3,00	♂♂	


a)
Taille, cm
Poids, g
2000
5000
57
47
45
43
41
b)
Taille, cm
Poids, g
1000
2000
4000
5000
57
49
47
45
43
41
39
37
35
33
c)
Taille, cm
Poids, g
1000
2000
4000
57
55
47
45
43
41
39
37
d)
Taille, cm
Poids, g
1000
2000
4000
5000
57
49
47
45
43
41
39
37
35

compliqué. Malgré la forte corrélation dans les groupes entre le poids et la taille du corps à la naissance (v. *Tableau 1*), la part des individus entrant dans le domaine « de norme adaptative » pour les deux caractères sera de beaucoup inférieure à celle présentée pour un seul caractère. Le principe de la classification des enfants d'après deux caractères quantitatifs est représenté sur les figures 3 et 4 sous forme d'une figure tridimensionnelle. Pour plus d'évidence, les carrés de même densité sont réunis sur la grille de corrélation par des isolignes par analogie avec une carte topographique. Sur le champ de corrélation on distingue 3 domaines principaux: central (M^0) correspondant au domaine de « norme adaptative » pour deux caractères; celui comprenant les enfants à grands indices poids-taille (M^+) et celui embrassant les enfants à bas indices poids-taille (M^-). On isole en outre 6 domaines complémentaires se trouvant à la périphérie de l'ellipse de corrélation (M^+/M^0, M^-/M^0, M^0/M^+, M^0/M^-, M^-/M^+, M^+/M^-) *. Ce principe de la classification peut être appliqué au dépouillement des données sur computer et pour le cas de n caractères.

Même à l'examen *de visu* des champs de corrélation, les différences entre les 5 groupes isolés d'enfants apparaissent avec une grande netteté. Chez les enfants sains, la distribution dans un espace tridimensionnel a la forme d'un sommet compact avec de petits contreforts. et dans un espace bidimensionnel, celle d'une ellipse dont les axes sont dans le rapport de 2 : 1 (v. fig. 3). Chez les enfants décédés, la distribution rappelle respectivement une chaîne fortement découpée par des contreforts et une ellipse dont les axes sont dans le rapport de 4 : 1 (v. fig. 4, *d*). La même asymétrie et la même déformation sont typiques des distributions des caractères anthropométriques dans les groupes de nouveau-nés malades observés en hôpital, de nourrissons malades et de nouveau-nés malades en maternité (fig. 4, *a*, *b*, *c*). Dans ce dernier cas, la distribution prend l'aspect d'une chaîne à deux sommets réunis par une petite barrette (fig. 4, *c*). En outre, il y a une déviation des distributions par rapport aux axes de coordonnées qui a pour conséquence une diminution de la part des individus à l'intérieur du « domaine adaptatif » parmi les enfants malades et décédés (*Tableau 2*). Chez les nouveau-nés sains, 37 % des enfants en font partie et parmi les décédés 13 %

* Dans chaque cas, la première désignation se rapporte à la taille; la seconde, au poids.

Fig. 4. Distribution conjointe du poids et de la taille à la naissance dans un groupe de nouveau-nés malades (*a*) et de nourrissons malades (*b*) observés en hôpital, de nouveau-nés malades observés en maternité (*c*) et dans un groupe d'enfants décédés en hôpital (*d*)

Le principe de construction et les désignation sont les mêmes que sur la figure 3

seulement (v. *Tableau 2*). En même temps, nous voyons que les enfants malades et décédés se trouvent plus souvent dans le domaine M^- de la grille de corrélation.

DISCUSSION

Nous avons montré les différences certaines de tous les paramètres statistiques des distributions de deux caractères anthropométriques des nouveau-nés sains et malades. Dans ce dernier cas, on observe une diminution de la part des sujets à valeurs moyennes des caractères et une augmentation de la fréquence des variantes phénotypiques « extrêmes ». En cette connexion, un problème se dresse: celui de l'identification du phénotype « moyen » d'après l'ensemble des caractères. Nous n'avons utilisé dans ce travail que deux caractères fortement corrélés (le poids et la taille à la naissance). Il est évident que si nous avions eu affaire à des caractères indépendants l'un de l'autre, à mesure de la croissance de leur nombre la part des phénotypes « moyens » dans l'échantillon étudié diminuerait, étant donné que dans un espace multidimensionnel un type « moyen » n'est qu'une abstraction. Aussi, lors de l'établissement d'une échelle unique pour identifier le phénotype « moyen », l'approche mathématique purement formelle n'a pas de sens et la préférence doit être accordée à l'approche biologique. En d'autres termes, il s'agit d'une différence qualitative entre les moyennes biologique et mathématique. S'il n'y a dans bien des cas derrière une moyenne mathématique que quelque chose de conventionnel, l'individu biologiquement « moyen » est une notion tout à fait réelle dans la mesure où la sélection naturelle impose des limites à la combinaison libre des caractères et établit des corrélations morphogénétiques rigides en rapport avec l'homéostasie ontogénétique [3, 4].

Par conséquent, l'approche des caractères choisis pour la classification ne peut pas être arbitraire: il faut accorder la préférence à un ensemble de caractères rattachés au nombre le plus grand possible de systèmes morphofonctionnels de l'organisme et caractérisés par un coefficient d'héritabilité assez élevé. Il est également évident que la tendance à l'élargissement illimité d'un jeu de caractères quantitatifs peut ne pas avoir de sens biologique.

On conçoit que l'abord qualitatif présenté ici n'est qu'une première approche du problème et qu'un travail ultérieur sérieux est nécessaire dans cette direction avec utilisation des méthodes de statistique multidimensionnelle. Cependant, déjà au niveau d'analyse atteint, l'adaptabilité supérieure des enfants « moyens » par le poids et la taille est tout à fait évidente d'après les matériaux envisagés ici. En lui-même, ce fait n'a rien d'étonnant. Depuis longtemps, la pratique médicale nous a montré que parmi les groupes d'enfants ayant des indices poids et taille hauts et, surtout, bas, la mortalité et la morbidité sont plus élevées qu'en moyenne dans la population

(v. par exemple [2]). En termes de génétique, cette causalité peut être interprétée comme une sélection stabilisante d'après les caractères quantitatifs aux stades ontogénétiques précoces. Cependant, il est nécessaire de savoir dans quelle mesure ce principe conserve son importance pour les stades plus avancés de l'ontogénèse. Peut-on extrapoler les résultats d'une telle classification des nouveau-nés pour les étapes plus avancées de l'ontogénèse postnatale, ainsi que pour les adultes?

Les données obtenues attestent que le schéma de classification proposé conserve son importance tout au moins jusqu'à 1 an d'âge (v. fig. 1-4). Apparemment, le même tableau peut être observé aux stades ontogénétiques ultérieurs. Cependant, la classification suivant les caractères anthropométriques sera d'autant plus difficile que l'influence du milieu sur leur expression phénotypique sera plus forte. C'est ce qui confère une importance particulière à l'analyse de la liaison entre les caractères quantitatifs et les caractères qui ont une base héréditaire claire et simple, considérés non pas isolément, mais au niveau d'un échantillon suffisamment représentatif de gènes du génome. Les premiers pas dans cette direction ont déjà été effectués [1]. La tâche consiste à étendre cette approche au plus grand nombre possible de gènes de localisation et de fonction différentes.

Mais même l'étude simultanée de 5 loci seulement de groupes sanguins met en évidence, nous l'avons vu plus haut, des différences essentielles dans la structure génétique entre les enfants morphologiquement « moyens », d'une part, et les groupes phénotypiques « extrêmes », de l'autre. Si cette même connexion se confirme dans les recherches ultérieures, une possibilité inconnue jusqu'alors s'offrira de prévoir l'« itinéraire individuel » de la santé dès la naissance et permettra d'aborder d'une façon nouvelle les problèmes de la prévention des maladies et de la planification de la Santé publique.

En ce qui concerne le groupe d'enfants que nous avons étudié, sa subdivision naturelle au moins en 2 sous-groupes est évidente: le premier sous-groupe embrasse les enfants dont les caractères anthropométriques ne sortent pas, par leurs variations, des limites de la « norme adaptative »; le second inclut les enfants s'écartant du centre de distribution de ces caractères vers la droite et, surtout, vers la gauche. L'analyse de la structure génétique des sous-groupes envisagés aura lieu dans les travaux suivants de ce cycle.

RÉFÉRENCES

[1] *Дубинин Н. П., Алтухов Ю. П., Курбатова О. Л., Сусков И. И.* Интегральная генетическая характеристика «адаптивной нормы» в популяциях человека. *Докл. АН СССР*, 1976, 230, 4, 957.

[2] *Karn M. N., Penrose L. S.* Birth weight and gestation time in relation to maternal age, parity and infant survival. *Ann. Eugenics*, 1951, 15, 206.

[3] *Lerner M. I.* Genetic gomeostasis. N.Y., Wiley, 1954.

[4] *Waddington C.* New patterns in genetics and development. N.Y., 1962.

APPROCHE DU PROBLÈME DE LA RÉSISTANCE BIOLOGIQUE NON SPÉCIFIQUE DE L'ORGANISME HUMAIN AU NIVEAU DE LA GÉNÉTIQUE DES POPULATIONS

Communication II. Caractéristique clinique, malformations et structure génétique des enfants malades compte tenu de leur poids et de leur taille à la naissance

O. Botviniev, O. Kourbatova et Y. Altoukhov

Nous avons démontré dans la communication précédente la liaison existant entre les caractères anthropométriques, poids et taille à la naissance, et la résistance aux maladies des enfants dans la première année de leur vie (v. pp. 208-217). Nous avons vu qu'alors que la variabilité de ces caractères dans le groupe d'enfants sains suit une loi normale, la distribution du poids et de la taille dans les groupes d'enfants malades et décédés est caractérisée par une moindre moyenne, une plus grande variance, une asymétrie à droite et par d'autres paramètres statistiques reflétant des troubles plus ou moins prononcés de l'ontogénèse. Ce sont les enfants entrant par deux caractères à la fois dans le domaine que nous distinguons comme « norme adaptative », c'est-à-dire le type morphologiquement « moyen », qui résistent le mieux à des maladies de différente nature. La part des individus faisant partie du domaine « de norme adaptative » se réduit en conséquence dans une dépendance directe par rapport à la charge pathologique totale du groupe d'enfants étudiés. La part des individus présentant des indices poids-taille extrêmes, surtout faibles, est particulièrement élevée parmi les enfants décédés.

En interprétant les causalités mises en évidence comme une manifestation de la sélection stabilisante d'après un système polygénique coadapté en accord avec les principes énoncés dans l'article précédent, nous allons exposer ici une information plus détaillée sur la structure de la morbidité dans les groupes étudiés et nous examinerons leur caractéristique génétique pour trois loci polymorphes. Ceci élargira nos connaissances du rôle de l'hérédité polygénique et de certains gènes mendéliens dans la résistance à diverses maladies, résistance qui détermine pour beaucoup la variabilité de l'organisme humain aux stades précoces de l'ontogénèse.

© Издательство Наука, 1980.

MATÉRIEL ET MÉTHODE

En accord avec le programme exposé plus haut (v. p. 208), 2160 enfants malades ont été observés dans la première année de leur vie (nouveau-nés en maternité et en hôpital, nourrissons, enfants morts). L'analyse de la morbidité a été effectuée compte tenu des maladies principale et concomitantes et en cas de mort, des données de l'autopsie. On a tenu compte des « grandes » tares et des « petites » tares dites stigmates.

Les groupes sanguins ABO et Rhésus ont été déterminés par les méthodes sérologiques standard chez 1874 enfants malades et décédés. L'otoscopie selon les critères de Matsunaga [6] a permis de définir les types de cérumen chez 493 malades. En outre, les groupes sanguins et les types de cérumen ont été déterminés chez des enfants pratiquement sains pris pour témoins.

La classification des enfants suivant le poids et la taille à la naissance a été faite selon les principes exposés dans la communication précédente: le groupe M^0 défini comme domaine « de norme adaptative » comprenait les enfants ayant des indices poids-taille moyens (taille 50 à 52 cm, poids 3200 à 3650 pour ♀ et 3250 à 3700 pour ♂). Les groupes M^- et M^+ embrassaient les enfants ayant respectivement des indices poids-taille bas et élevés. Le groupe M^0/M^- incluait les enfants de taille moyenne et de poids inférieur à la normale, le groupe M_{pd} (« phenodeviants ») les autres enfants présentant une corrélation perturbée entre le poids et la taille à la naissance * (M^0/M^+, M^-/M^0, M^+/M^0, M^-/M^+, M^+/M^-).

Les diverses formes de morbidité et les particularités génétiques des enfants du premier âge sont examinées sur le plan de la comparaison des enfants groupés d'après le poids et la taille à la naissance. De plus, on a étudié aussi bien le niveau général de la morbidité dans les groupes que sa structure spécifique.

PARTIE EXPÉRIMENTALE

Morbidité. La structure de la morbidité des groupes d'enfants étudiés se caractérise par un nombre varié d'entités nosologiques: 32 chez les nouveau-nés et les nourrissons, 47 chez les décédés. Le *Tableau 1* résume les diagnostics le plus souvent constatés (v. maladies nos 1-16). Les maladies génétiques (mucoviscidose, ichtyose, syndrome génito-surrénal, anémie du type Minkowski-Chauffard) forment un groupe spécial (v. *Tableau 1*, n° 1). Les différentes maladies, abstraction faite de celles de numéros 1-16, sont rangées en catégorie « autres » (v. *Tableau 1*, n° 17). Remarquons qu'en comparaison des nourrissons, les nouveau-nés sont plus souvent sujets aux

* Dans chaque cas, la première désignation se rapporte à la taille; la seconde au poids.

Affections observées chez les enfants malades et décédés classés suivant le

Maladie		Nouveau-nés malades					
Numéro	Nom	M^0	M^-	M^+	M^0/M^-	M_{pd}	Σ
1	Maladies génétiques	1	1	0	2	5	1
2	Affection respiratoire aiguë	36	32	29	37	36	34
3	Septicémie	10	25	7	7	7	12
4	Pneumonie	18	29	20	30	25	24
5	Pyodermite	40	25	40	38	38	36
6	Conjonctivite	17	13	17	23	23	18
7	Otite	33	29	27	27	24	29
8	Muguet	28	20	25	27	31	26
9	Diarrhée	13	15	13	14	19	14
10	Encéphalopathie posthypoxémique	61	70	64	65	59	64
11	Traumatisme intracrânien obstétrical	4	6	7	4	5	5
12	Hypotrophie	12	19	5	23	8	14
13	Anémie	1	7	2	9	3	4
14	Diathèse exsudative	15	8	18	14	16	14
15	Ictère des nouveau-nés	19	36	19	32	24	26
16	Crise génitale	19	7	16	17	22	16
17	Autres	29	28	45	43	52	37
Nombre total des enfants		221	163	118	124	87	713
Nombre de maladies par enfant		3,58± ±0,13	3,69± ±0,15	3,53± ±0,17	4,14± ±0,18	3,96± ±0,21	3,74± ±0,07

infections pyogènes et mycosiques : ils sont plus fréquemment atteints de pyodermite, de septicémie, de muguet, de conjonctivite (v. *Tableau 1*). En outre, chez les nouveau-nés on observe plus fréquemment des troubles du SNC. Chez les nourrissons, ce sont les affections aiguës des organes respiratoires, les pneumonies, les anémies, l'hypotrophie. Le groupe d'enfants décédés se caractérise par la présence de pneumonies, de septicémie, d'atteintes organiques du SNC, d'ictère, de maladies génétiques. Les différentes affections rangées en catégorie « autres » (n° 17) sont surtout observées chez les nouveau-nés.

Lorsqu'on compare la structure de la morbidité parmi les enfants groupés d'après leur poids et leur taille à la naissance (v. *Tableau 1*), on voit qu'il existe certaines maladies distribuées régulièrement entre toutes les catégories d'enfants. Ce sont les affections respiratoires aiguës, la conjonctivite, l'otite, le muguet, le syndrome intestinal, le traumatisme intracrânien obstétrical. En même temps, il y a des différences nettes quant à certaines maladies dont beaucoup peuvent évoluer vers une issue fatale. Ainsi, les maladies géné-

Tableau 1

poid et la taille à la naissance (pour 100 enfants dans un groupe étudié)

Nourrissons malades						Enfants décédés					
M^0	M^-	M^+	M^0/M^-	M_{pd}	$\sum$	M^0	M^-	M^+	M^0/M^-	M_{pd}	$\sum$
1	1	1	1	3	1	5	3	6	3	14	4
76	79	85	77	81	79	9	8	30	6	14	10
2	7	5	3	3	5	31	33	26	54	31	35
27	56	42	33	37	41	64	79	82	72	91	76
22	14	33	24	22	21	3	8	18	7	17	8
4	8	2	7	2	5						
37	31	40	33	37	35	12	5	10	7	9	7
6	13	8	8	5	9						
13	13	12	13	11	13	14	19	18	10	11	16
34	40	35	45	28	37	15	22	12	14	14	19
1	8	4	1	0	4	35	29	32	30	31	31
22	27	13	20	14	21	10	6	18	6	14	9
19	34	19	22	8	24	8	6	2	7	3	6
60	48	67	56	55	56						
						52	85	36	61	49	70
21	17	21	26	9	19	25	30	26	30	29	28
168	222	91	89	64	634	127	337	50	69	35	618
,48±	3,96±	3,87±	3,70±	3,16±	3,70±	2,83±	3,35±	3,16±	3,09±	3,29±	3,19±
±0,14	±0,13	±0,21	±0,20	±0,22	±0,08	±0,15	±0,10	±0,25	±0,21	±0,31	±0,07

tiques intéressent plus souvent les enfants M_{pd}, aussi bien nouveau-nés que nourrissons et décédés. La septicémie est surtout fréquente chez les enfants M^- (nouveau-nés, nourrissons) et chez les enfants M^0/M^- (morts). Les enfants M^0 contractent moins souvent la pneumonie et les enfants M^-, plus souvent. En outre, on découvre aussi les différences suivantes: la pyodermite est plus fréquente chez les enfants M^+ (nourrissons et décédés); les troubles du SNC. les anémies, l'hypotrophie prévalent chez les enfants M^-; l'ictère des nouveau-nés intéresse surtout les enfants M^- et M^0/M^-. Cependant, la crise génitale est beaucoup plus rare chez les enfants M^- que chez les enfants M^+, ainsi que la diathèse exsudative. La catégorie « autres » est surtout typique des enfants M_{pd}. Comme le montrent le *Tableau 1* et l'étude globale de toutes les maladies, leur nombre pour un enfant M^0 est toujours moindre que pour un enfant M^- ou M^0/M^-.

Ainsi, l'étude clinique atteste que le niveau et la structure de la morbidité parmi les enfants à indices poids-taille différents diffèrent aussi. Il est évident que les enfants entrant dans le domaine de «norme adaptative» (M^0) tombent moins souvent malades et leurs affections sont plus bénignes (fig. 1). Il y a lieu de citer ici que la part

« Grandes » et « petites » tares (stigmates) chez les enfants mala (100 enfants dans

Malformations		Nouveau-nés malades					
Numéro	Désignation	M^0	M^-	M^+	M^0/M^-	M_{pd}	Σ
1	Malformations multiples	2	7	0	1	4	3
2	Cardiopathies congénitales	0	10	2	4	2	3
3	Malformations du système nerveux						
4	Malformations du tube digestif						
5	Malformations de l'appareil urinaire						
6	Malformations de l'appareil génital	10	18	9	15	4	12
7	Malformations ostéomusculaires	2	12	6	7	10	7
8	Malformations crâniennes	9	13	17	15	10	13
9	Hernies inguino-scrotale et ombilicale	5	13	5	3	4	6
10	Autres malformations						
11	Anomalies congénitales du pavillon de l'oreille	15	27	21	24	14	20
12	Palais ogival	5	8	6	4	8	6
13	Télangiectasie	29	20	16	15	26	22
14	Anomalies congénitales de l'appendice xiphoïde	5	12	9	12	4	8
15	Anomalies congénitales de la région coccygienne	25	25	22	33	18	25
16	Plis palmaires anormaux	8	8	8	4	10	7
17	Thymomégalie *	11	10	13	4	10	10
18	Autres stigmates	18	23	16	12	10	16
	Total des enfants	118	60	63	72	50	363
	Nombre de malformations par enfant	1,45± ±0,11	2,06± ±0,19	1,51± ±0,15	1,56± ±0,15	1,34± ±0,16	1,57± ±0,07

* L'augmentation du thymus n'est pas considérée comme une tare, mais comme une

des enfants M^0 parmi les décédés et les enfants malades n'est que respectivement de 13 % et de 19 à 28 %, alors que parmi les enfants sains elle atteint 37 % [1]. Les enfants M^- et M^0/M^- tombent plus souvent et plus gravement malades. Les enfants M^+ et M_{pd} occupent une position intermédiaire.

Malformations. Le *Tableau 2* montre les données sur les tares « grandes » et « petites » chez les enfants malades et décédés. Tous les groupes étudiés n'ont pas été caractérisés exhaustivement sous ce

Tableau 2

des et décédés classés suivant le poids et la taille à la naissance un groupe étudié)

Nourrissons malades						Enfants décédés					
M^0	M^-	M^+	M^0/M^-	M_{pd}	Σ	M^0	M^-	M^+	M^0/M^-	M_{pd}	Σ
1	3	2	0	2	2	15	22	14	16	14	19
2	11	5	5	2	6	16	9	24	15	6	12
						6	3	0	4	6	4
						21	20	10	32	14	20
						4	3	2	1	6	3
10	13	11	16	12	12	2	2	10	3	3	3
10	13	5	18	12	12	3	1	2	4	6	2
20	27	12	20	19	21						
9	15	4	11	5	10						
						6	7	6	16	20	9
10	17	11	13		13						
17	22	23	20	17	20						
24	17	18	31	27	21						
16	15	11	9	12	14						
20	13	7	7	19	14						
10	11	12	9	5	10						
13	14	37	18	22	18						
29	35	30	49	24	34						
97	151	56	55	41	400	127	337	50	69	35	618
1,92±	2,26±	1,89±	2,27±	1,93±	2,09±	0,74±	0,68±	0,68±	0,91±	0,74±	0,72±
±0,14	±0,12	±0,18	±0,20	±0,22	±0,07	±0,08	±0,04	±0,12	±0,11	±0,15	±0,03

particularité phénotypique.

rapport. Ainsi, on n'a pas toujours remarqué les « petites » tares (stigmates) parmi les décédés, c'est pourquoi on n'en a pas tenu compte dans l'étude. On a relevé en tout 39 types de malformations chez les nouveau-nés, 51 chez les nourrissons et 56 chez les enfants morts. Dans le *Tableau 2*, elles sont rangées en 18 catégories.

Parmi les « grandes » tares chez les nouveau-nés et les nourrissons, on a le plus souvent signalé des malformations crâniennes, des anomalies congénitales de l'appareil génital et du système ostéomus-

Groupes sanguins dans quelques groupes de nouveau-nés et de

Système	Paramètres génétiques	Nouveau-nés sains	Nouveau-nés		
			M^0	M^-	M^+
ABO	p	0,267±0,015	0,257±0,024	0,216±0,025	0,267±0,032
	q	0,154±0,011	0,150±0,019	0,131±0,020	0,192±0,027
	r	0,579±0,017	0,593±0,027	0,653±0,030	0,541±0,036
	h	0,569±0,011	0,584±0,019	0,498±0,026	0,558±0,021
	χ^2	0,007	2,491	0,562	3,202
Rhésus	d	0,419±0,020	0,417±0,033	0,326±0,039	0,432±0,042
	h	0,487±0,005	0,486±0,009	0,439±0,019	0,491±0,009
	N	527	196	151	114

Système	Paramètres génétiques	Nourrissons malades			
		M^0/M^-	M_{pd}	$\sum$	M^0
ABO	p	0,223±0,034	0,243±0,042	0,247±0,013	0,178±0,026
	q	0,186±0,032	0,141±0,033	0,155±0,011	0,193±0,026
	r	0,592±0,040	0,616±0,047	0,60±0,015	0,630±0,033
	h	0,565±0,029	0,539±0,037	0,563±0,012	0,523±0,027
	χ^2	0,003	0,012	0,879	0,393
Rhésus	d	0,453±0,048	0,471±0,057	0,415±0,018	0,363±0,042
	h	0,496±0,008	0,498±0,008	0,486±0,005	0,462±0,017
	N	86	61	620	124

Remarque. h est la valeur de l'hétérozygotie pour un locus donné.
Les astérisques indiquent respectivement les valeurs certaines au niveau
* $P > 0,95$,
** $P > 0,99$,
*** $P > 0,999$.

culaire; parmi les « petites » tares chez les nouveau-nés on a constaté des anomalies congénitales du coccyx et du pavillon de l'oreille, une télangiectasie. Chez les nourrissons, on voit prédominer les autres stigmates, puis la télangiectasie et le palais ogival. En outre, chez les nourrissons la radiographie frontale révèle assez souvent le thymus. La ressemblance de la structure de toutes les malformations étant très grande, le niveau général de ces anomalies est plus élevé chez les nourrissons que chez les nouveau-nés. Dans le groupe d'enfants décédés, les malformations du tube digestif occupent la première place; les malformations multiples, la seconde et les cardiopathies

Tableau 3

nourrissons classés suivant le poids et la taille à la naissance

malades			Nourrissons malades		
M^0/M^-	M_{pd}	$\sum$	M^0	M^-	M^+
0,266±0,031	0,246±0,037	0,249±0,013	0,239±0,025	0,254±0,022	0,249±0,035
0,169±0,026	0,139±0,028	0,154±0,010	0,153±0,021	0,162±0,018	0,125±0,026
0,565±0,035	0,615±0,041	0,597±0,014	0,609±0,029	0,584±0,026	0,626±0,039
0,592±0,022	0,548±0,032	0,560±0,011	0,579±0,022	0,578±0,018	0,513±0,031
0,264	0,059	0,025	3,221	0,394	0,707
0,443±0,041	0,314±0,053	0,394±0,018	0,401±0,036	0,397±0,031	0,403±0,048
0,494±0,008	0,431±0,027	0,478±0,006	0,481±0,011	0,479±0,010	0,486±0,015
119	81	678	165	218	90

Enfants décédés

M^-	M^+	M^0/M^-	M_{pd}	$\sum$
0,230±0,020	0,213±0,047	0,202±0,038	0,270±0,058	0,216±0,013
0,152±0,016	0,155±0,040	0,117±0,030	0,159±0,046	0,157±0,011
0,618±0,023	0,632±0,056	0,681±0,044	0,572±0,064	0,627±0,015
0,501±0,018	0,465±0,046	0,498±0,042	0,518±0,042	0,504±0,007
11,440 ***	4,254 *	0,484	2,166	12,790 ***
0,400±0,028	0,363±0,071	0,389±0,058	0,455±0,075	0,391±0,019
0,480±0,009	0,462±0,029	0,475±0,020	0,496±0,015	0,477±0,007
262	43	63	35	576

congénitales, la troisième. La charge globale des tares « graves » chez les enfants décédés est beaucoup plus lourde que parmi les nouveau-nés et nourrissons malades (v. *Tableau 2*).

La comparaison de la structure des malformations dans les groupes d'enfants classés d'après le poids et la taille montre que les enfants M^0 et M^+ ont moins de « grandes » et « petites » tares que les enfants M^- et M^0/M^- (v. *Tableau 2*, fig. 2). Il ressort en même temps un complexe de malformations typique des enfants ayant un poids et une taille déterminés à la naissance. Ce qui caractérise surtout les enfants M^-, c'est une grande fréquence de malformations multiples,

Distribution des phénotypes de deux loci des groupes sanguins suivant le poids et la

Numéro du phénotype	Phénotype	Fréquence théorique	Enfants sains	Nouveau-nés malades					
				M^0	M^-	M^+	M^0/M^-	M_{pd}	$\sum$
1	ARh^+	0,314	0,312	0,292	0,305	0,355	0,250	0,366	0,307
2	ORh^+	0,276	0,269	0,282	0,344	0,215	0,295	0,310	0,291
3	BRh^+	0,167	0,175	0,164	0,205	0,215	0,179	0,183	0,187
4	$ABRh^+$	0,068	0,068	0,087	0,040	0,028	0,080	0,042	0,060
5	ARh^-	0,067	0,068	0,067	0,046	0,056	0,098	0,014	0,060
6	ORh^-	0,059	0,066	0,067	0,060	0,028	0,045	0,056	0,054
7	BRh^-	0,036	0,025	0,021	0	0,075	0,027	0	0,024
8	$ABRh^-$	0,015	0,016	0,021	0	0,028	0,027	0,028	0,019
		χ^2	1,29	1,41	11,58 *	14,13 **	4,06	7,06	2,36

Remarque. Les désignations sont les mêmes que dans le *Tableau 3*.

Types de cérumen chez les enfants sains et malades

Paramètres génétiques		Enfants sains	Enfants	
			M^0	M^-
Type de cérumen				
sec ww	NO	198	76	98
	FO	0,377	0,551	0,700
liquide WW+	NO	327	62	42
+Ww	FO	0,623	0,449	0,300
Fréquence du gène w		0,614±0,017	0,742±0,028	0,837±0,023
h		0,474±0,007	0,383±0,025	0,273±0,030
N		525	138	140

Remarque. NO est le nombre observé d'enfants ayant ce phénotype ; FO la fréquence
Les autres désignations sont les mêmes que pour les *Tableaux 1-4*.

Tableau 4

parmi plusieurs groupes de nouveau-nés et de nourrissons classés taille à la naissance

Nourrissons malades						Enfants morts					
M^0	M^-	M^+	M^0/M^-	M_{pd}	$\sum$	M^0	M^-	M^+	M^0/M^-	M_{pd}	$\sum$
0,262	0,305	0,375	0,241	0,333	0,297	0,211	0,324	0,342	0,283	0,379	0,295
0,342	0,284	0,275	0,301	0,278	0,306	0,369	0,275	0,263	0,359	0,241	0,299
0,141	0,188	0,138	0,181	0,130	0,162	0,246	0,214	0,263	0,151	0,172	0,213
0,094	0,066	0,050	0,072	0,037	0,069	0,044	0,027	0	0,057	0	0,030
0,074	0,046	0,025	0,072	0,037	0,053	0,053	0,053	0,053	0,038	0	0,048
0,020	0,076	0,088	0,048	0,074	0,059	0,018	0,065	0,053	0,113	0,069	0,058
0,054	0,015	0,050	0,072	0,093	0,046	0,044	0,042	0,026	0	0,103	0,052
0,013	0,020	0	0,012	0,019	0,014	0,018	0	0	0	0,035	0,006
12,40 *	2,48	7,04	6,87	9,17	5,92	15,13 **	13,21 *	5,02	6,13	10,70 *	16,36 **

Tableau 5

classés suivant le poids et la taille à la naissance

malades			
M^+	M^0/M^-	M_{pd}	$\sum$
51	58	28	311
0,622	0,699	0,560	0,631
31	25	22	182
0,378	0,301	0,440	0,369
0,789±0,034	0,836±0,030	0,748±0,047	0,794±0,014
0,333±0,037	0,274±0,038	0,377±0,043	0,327±0,015
82	83	50	493

observée des enfants ayant ce phénotype.

de malformations crâniennes, de cardiopathies congénitales, ainsi que la présence de hernies, d'anomalies congénitales du pavillon de l'oreille et de l'appendice xiphoïde, d'un palais ogival. Chez les enfants M^+, la fréquence des cardiopathies congénitales est aussi élevée, assez souvent on constate chez eux des anomalies congénitales du système génital. Remarquons qu'aucun enfant à grands poids et taille à la naissance ne présentait de malformations du système nerveux. Les enfants M^0/M^- sont caractérisés par une fréquence élevée de malformations multiples, en particulier d'anomalies congénitales du tube digestif, des os et des muscles. Les anomalies ostéomusculaires et « autres tares» sont surtout observées chez les enfants M_{pd} (v. fig. 2).

Structure génétique. Le *Tableau 3* présente les fréquences géniques de groupes sanguins des systèmes ABO et Rhésus chez les enfants sains, malades (nouveau-nés et nourrissons) et décédés classés suivant leur poids et taille à la naissance. Pour tous les groupes que nous avons étudiés, on a calculé la valeur de l'hétérozygotie pour chacun de ces loci, ainsi que le degré de concordance de la distribution empirique des phénotypes avec la condition d'équilibre de Hardy-Weinberg (χ^2).

Etant donné que chez les nouveau-nés sains, on n'a pas découvert de différences dans les fréquences des gènes A, B, O et d, ainsi que dans les valeurs de l'hétérozygotie entre les groupes M^0, M^-, M^+, M^0/M^- et M_{pd}, les nouveau-nés sains sont représentés globalement dans le *Tableau 3*. En outre, la valeur minimale de χ^2 témoigne d'un accord parfait entre la distribution observée des phénotypes dans ce groupe et celle calculée en vertu de la formule de Hardy-Weinberg. Simultanément, ce qui est typique des enfants malades et décédés, c'est la croissance de la fréquence du gène O dans l'ordre suivant: nouveau-nés, nourrissons, enfants morts. Dans le même sens, on observe une diminution de la fréquence du gène A et une augmentation de la valeur de χ^2. Pour le système Rhésus on n'a pas constaté de différences significatives.

Dans le groupe de nouveau-nés malades, la fréquence du gène O est particulièrement prononcée parmi les enfants M^-. L'hétérozygotie pour le système ABO chez les nouveau-nés malades ne diffère pas de celle des nouveau-nés sains. Cependant, elle est notablement plus basse chez les enfants M^-, aussi bien pour le système ABO que pour le système Rhésus. Chez les nourrissons malades, la fréquence du gène O est maximale et l'hétérozygotie est minimale chez les enfants M^+. Dans le groupe d'enfants morts, la fréquence du gène O est également élevée chez les enfants M^+ et, en particulier, chez les enfants M^0/M^-. Ils sont hétérozygotes au minimum. Pour l'ensemble, dans le groupe d'enfants morts, l'hétérozygotie, tant pour le système ABO que pour le système Rhésus, est d'une façon certaine inférieure à celle des enfants sains. En outre, chez les enfants décé-

dés, on note un écart important de l'équilibre Hardy-Weinberg, surtout pour les enfants M^- et M^+.

Ainsi, le groupe d'enfants le plus taré est caractérisé par la plus grande fréquence du gène O, par un plus grand degré d'homozygotie et un écart significatif des fréquences des phénotypes de l'équilibre

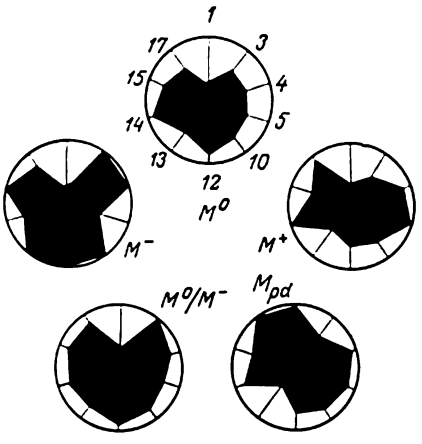


Fig. 1. Fréquence relative de certaines maladies dans cinq groupes d'enfants distingués par le poids et la taille du corps à la naissance (total des nouveau-nés, des nourrissons malades et des morts).

Sur chaque rayon on a porté la fréquence relative d'une maladie dans un groupe donné : zéro au centre du cercle, 100 % sur le pourtour. On a pris pour 100 % la fréquence maximale d'une maladie donnée dans les groupes. Les maladies sont numérotées de la même façon qu'au *Tableau 1*

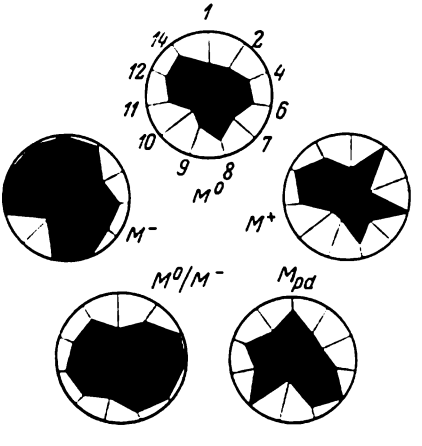


Fig. 2. Fréquence relative des malformations dans cinq groupes d'enfants distingués par le poids et la taille du corps à la naissance (total des nouveau-nés, des nourrissons malades et des morts).

Le principe de construction est le même que sur la figure 1. Les malformations sont numérotées de la même façon qu'au *Tableau 2*

Hardy-Weinberg. Afin d'élucider le rôle joué dans cette sélection par l'appartenance Rhésus, le *Tableau 4* montre les combinaisons phénotypiques de deux loci dans l'ordre de décroissance de leur fréquence théorique dans la population. Selon le critère χ^2, la distribution empirique est, chez les enfants sains, en bon accord avec la distribution théorique. En même temps, dans les groupes d'enfants malades et surtout décédés, l'écart de la distribution théorique devient significatif. Dans certains groupes, l'ordre théorique des fréquences phénotypiques est perturbé. Par exemple, chez les enfants M^0 morts, le phénotype ARh^+, le plus fréquent théoriquement dans la population, est observé plus rarement que le deuxième ORh^+ et le troisième BRh^+. L'analyse montre que la fréquence du phénotype ORh^+ chez les enfants malades et décédés est partout plus élevée que chez

les enfants sains. Pour les autres phénotypes, une dépendance aussi nette n'est pas notée. Cependant, parmi les enfants morts on relève une augmentation de la fréquence BRh^+ et BRh^-.

Le *Tableau 5* présente les données qui se rapportent aussi à un marqueur génétique (le type de cérumen) et qui confirment la dépendance notée plus haut. Chez les enfants malades, la fréquence du gène récessif du cérumen sec w est plus élevée de façon certaine que chez les enfants sains, tandis que le niveau d'hétérozygotie est plus bas. En outre, les enfants M^0 sont notablement plus rapprochés des témoins que tous les autres en ce qui concerne la fréquence du gène w et le niveau d'hétérozygotie. La fréquence du gène du cérumen sec est particulièrement élevée et le niveau d'hétérozygotie bas chez les enfants M^- malades.

DISCUSSION

Les données envisagées montrent l'existence, chez les nouveau-nés et les nourrissons, d'un lien entre les caractères anthropométriques normaux et la résistance aux maladies, d'une part, et entre la résistance aux maladies et la structure génétique, de l'autre. Nous voyons clairement la distribution inégale de la morbidité parmi les enfants classés d'après le poids et la taille à la naissance. Pratiquement, toutes les maladies et malformations étudiées (surtout les graves) intéressent beaucoup plus souvent les enfants ayant des caractères anthropométriques de valeur extrême (v. fig. 1 et 2). On observe aussi une spécificité déterminée dans la structure de la morbidité des groupes comparés : si des maladies comme la septicémie, la pneumonie, l'hypotrophie, l'anémie, l'ictère des nouveau-nés, l'encéphalopathie posthypoxémique, les tares multiples, les hernies, les malformations crâniennes et certains stigmates frappent surtout les enfants M^-, la pyodermite, la diathèse exsudative, les malformations de l'appareil génital sont plus communes aux enfants M^+. Dans l'ensemble, il apparaît nettement la stabilité maximale des individus morphologiquement « moyens » envers une large gamme de maladies d'étiologie diverse. Les rapports de cause à effet qui sont à la base de ces dépendances peuvent être variés. Ainsi, la plupart des malformations que nous avons étudiées sont de nature génétique et il y a une corrélation déterminée entre le génotype anormal et le phénotype [1]. En ce sens, n'importe quel écart des paramètres anthropométriques peut résulter de troubles de l'embryogénèse imputables à un génotype anormal du fœtus et (ou) aux facteurs du milieu. Cependant, des dimensions du corps du nouveau-né s'écartant fortement de la moyenne de la population peuvent par elles-mêmes être un facteur prédisposant à la maladie (prématurité, gros enfant). En outre, on connaît de nombreux exemples d'associations entre les différentes maladies et les marqueurs génétiques. Toutes ces interactions diver-

ses entre le génotype et le milieu au cours de l'évolution devaient conduire inéluctablement à l'apparition d'un complexe coadapté de gènes résultant de la « canalisation » du développement. Ce système génétique intégré doit déterminer aussi bien les particularités morphologiques de l'organisme que sa résistance non spécifique aux maladies. Les théories de l'homéostasie génétique [3, 5, 7] rattachent le phénotype le plus adapté à l'hétérozygotie pour la totalité des loci, le phénotype non adapté (« phenodeviant ») étant signe qu'un certain seuil d'hétérozygotie n'a pas été atteint.

Pour la première fois, un complexe coadapté de gènes a été décrit pour la drosophile [3]. Naguère encore, la question de l'existence d'un tel complexe de gènes chez l'homme restait ouverte (v. [2] par exemple). Il est clair maintenant, que l'homme ne fait pas exception aux lois biologiques essentielles. Cependant, le caractère et la nature des associations génétiques correspondantes ne peuvent être révélés qu'en augmentant considérablement le lot de gènes étudié. Les résultats présentés dans cet article et ceux publiés antérieurement [4] montrent qu'au moins une conclusion reste inchangée: celle des différences dans la structure génétique des enfants classés, aux stades ontogénétiques précoces, suivant les caractères du développement physique. En ce qui concerne les loci géniques y envisagés, la différenciation est en rapport avec le niveau plus bas d'hétérozygotie chez les individus M^- qui sont en même temps les plus tarés. Il est possible que pour d'autres loci, cette dépendance soit différente. Mais pour le génome pris en entier sera juste, à notre point de vue, un modèle supposant une diminution de l'adaptabilité des associations génotypiques rares, indifféremment du caractère de la distribution de l'hétérozygotie. Actuellement, cette question est étudiée par les auteurs de cet article pour 7 loci des groupes sanguins et 18 loci contrôlant la synthèse des protéines enzymatiques et non enzymatiques.

RÉFÉRENCES

[1] *Бочков Н. П.* Генетика человека (наследственность и патология). М., Медицина, 1978.
[2] *Dobzhansky T. H.* Genetics of the evolutionary process. N.Y.-L., Columb. Univ. Press, 1970.
[3] *Дубинин Н. П.* Экспериментальное исследование интеграции наследственных систем в процессах эволюции популяций. *Ж. общ. Биол.*, 1948, 9, 3, 203.
[4] *Дубинин Н. П., Алтухов Ю. П., Курбатова О. Л., Сусков И. И.* Интегральная генетическая характеристика « адаптивной нормы » в популяциях человека. *Докл. АН СССР*, 1976, 230, 4, 957.
[5] *M. I. Lerner.* Genetic gomeostasis. N.Y., Wiley, 1954.
[6] *Matsunaga E.* The dimorphism in human normal cerumen. *Ann. Human. Genet. Lond.*, 1962, 25, 4, 273.
[7] *Waddington C.* New patterns in genetics and development. N.Y. 1962.

RÉSISTANCE NON SPÉCIFIQUE DE L'ORGANISME ET PRÉDISPOSITION À LA LEUCÉMIE AIGUË PARMI LES ENFANTS

O. Kourbatova, O. Botviniev, M. Dechtchokina, I. Torbiak et Y. Altoukhov

Les articles précédents (v. pp. 208 et 218) ont montré le lien existant entre les caractères anthropométriques à la naissance, les particularités génétiques et la résistance à une large gamme de maladies chez les nouveau-nés et les nourrissons. On a vu que les enfants présentant un écart certain de la moyenne de la population pour la totalité des caractères anthropométriques offrent moins de résistance non spécifique aux maladies et ont une structure génétique bien déterminée. Dans leur interprétation des lois mises en évidence dans le cadre de la conception de la « norme adaptative » [6], les auteurs ont avancé l'hypothèse de l'existence d'un complexe coadapté de gènes déterminant les particularités constitutionnelles de l'organisme humain et ses possibilités homéostatiques de résister à différentes fluctuations du milieu, tant intérieur qu'extérieur [2].

Dans le présent article, nous avons examiné le rôle de la résistance non spécifique générale dans la genèse de la leucémie aiguë chez les enfants.

MATÉRIEL ET MÉTHODE

Nous avons dépouillé 422 dossiers médicaux d'enfants atteints de leucémie aiguë (229 ♂♂ et 193 ♀♀) pris dans les archives de plusieurs hôpitaux de Moscou pour les années 1959-1979. On a étudié les caractères anthropométriques, poids et taille des enfants à la naissance, ainsi que les groupes sanguins ABO et Rhésus, l'âge au début de la leucémie, les maladies contagieuses et inflammatoires ayant précédé la leucémie, l'âge des parents à la naissance de l'enfant, le numéro d'ordre de la grossesse et de l'accouchement. Un lot d'enfants pratiquement sains (396 sujets) a servi de groupe témoin.

PARTIE EXPÉRIMENTALE

La figure 1 montre les distributions du poids et de la taille à la naissance dans un groupe de leucémiques et dans celui de témoins bien portants.

© Издательство Наука, 1982.

Tableau 1

Principaux paramètres statistiques des distributions des leucémiques aigus et des enfants sains d'après le poids et la taille à la naissance

Caractère	Paramètres statistiques	Enfants observés		P
		atteints de leucémie aiguë	sains	
Poids à la naissance, g	$M \pm m$	3491±27	3458±21	$< 0,95$ ($t_d = 0,96$)
	σ	515	428	$> 0,999$ ($F = 1,45$)
	$\gamma_A \pm m_A$	0,0629±0,1300	0,3125±0,1231	$< 0,95$ ($t_d = 1,39$)
	$\gamma_E \pm m_E$	3,9790±0,2600	3,2185±0,2462	$> 0,95$ ($t_d = 2,12$)
Taille à la naissance, cm	$M \pm m$	51,39±0,16	51,16±0,09	$< 0,95$ ($t_d = 1,25$)
	σ	2,67	1,89	$> 0,999$ ($F = 2,0$)
	$\gamma_A \pm m_A$	−0,0523±0,1477	0,2812±0,1228	$< 0,95$ ($t_d = 1,74$)
	$\gamma_E \pm m_E$	3,9280±0,2954	2,8829±0,2456	$> 0,99$ ($t_d = 2,72$)
	N	355	396	

Remarque. M est la moyenne du caractère, m son erreur ;
σ l'écart type ;
γ_A l'indice d'asymétrie, m_A son erreur ;
γ_E l'indice d'excès, m_E son erreur ;
t_d le critère de Student ;
F le critère de Fisher ;
N le nombre d'individus observés.

Tableau 2

Distribution des enfants leucémiques et sains par classes compte tenu des caractères anthropométriques

Enfants examinés		Classe des caractères anthropométriques										Critère de certitude des différences
		M^0	M^+	M^-	M^0/M^-	M^0/M^+	M^-/M^0	M^+/M^0	M^+/M^-	M^-/M^+	$\sum$	
Atteints de leucémie aiguë	N	84	60	40	34	27	6	20	3	1	275	$\chi^2 = 21,67$ $P > 0,99$
	%	30,55	21,82	14,55	12,36	9,82	2,18	7,27	1,09	0,36	100	
Sains	N	140	93	40	48	13	2	18	2	0	356	
	%	39,33	26,12	11,24	13,48	3,65	0,56	5,06	0,56	0	100	

Remarque. Les désignations sont les mêmes que sur la figure 2 et le *Tableau 1*.

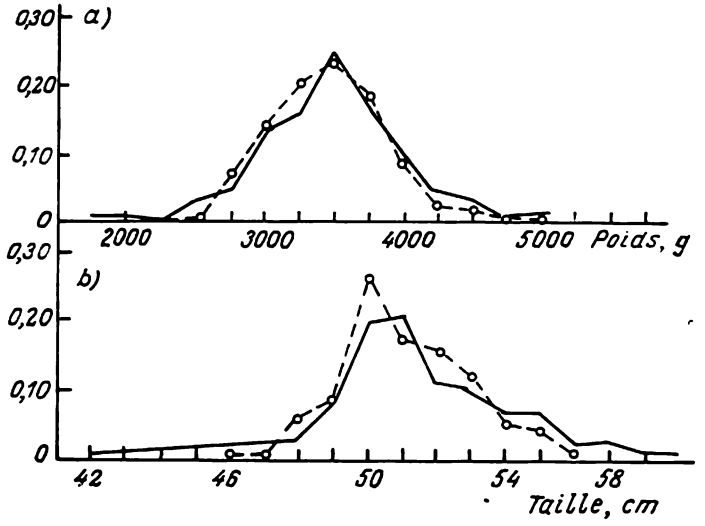


Fig. 1. Distribution conjointe du poids (*a*) et de la taille (*b*) à la naissance dans un groupe d'enfants atteints de leucémie aiguë (ligne continue) et dans le groupe témoin (ligne pointillée)

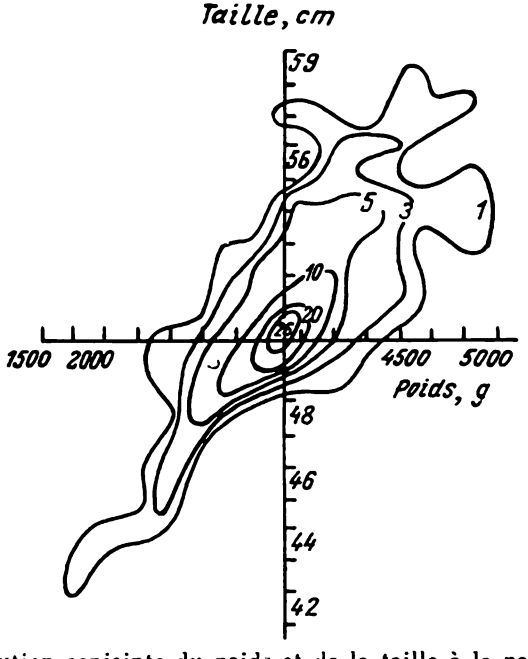


Fig. 2. Distribution conjointe du poids et de la taille à la naissance dans un groupe d'enfants atteints de leucémie aiguë:

Les chiffres sur les isolignes figurent le nombre d'enfants entrant dans la zone de grille de corrélation par le milieu de laquelle passe l'isoligne

Le *Tableau 1* présente les principaux paramètres statistiques des distributions du poids et de la taille. La comparaison des deux distributions montre que les moyennes étant presque égales, les variances des deux caractères dans le groupe d'enfants malades sont authentiquement plus élevées que chez les bien portants. Alors que chez les témoins, les distributions du poids et de la taille suivent une loi normale, dans le groupe de leucémiques les distributions sont déformées de sorte à engendrer des fréquences cumulées dans le domaine des valeurs extrêmes des caractères, ce qui entraîne un excès positif de la distribution (v. *Tableau 1*).

Tournons-nous maintenant vers la comparaison des distributions conjointes du poids et de la taille chez les enfants sains * et malades (fig. 2). Le principe de la construction de ces distributions et de la classification des individus suivant deux caractères quantitatifs a été décrit plus haut (v. p. 215).

La comparaison des distributions qui sont typiques des enfants sains (v. fig. 3, p. 212) et de ceux atteints de leucémie aiguë (v. fig. 2), révèle des différences évidentes entre ces deux groupes. Chez les enfants sains, la distribution dans un espace bidimensionnel a la forme d'une ellipse dont les axes sont dans le rapport de 2 : 1, et dans un espace tridimensionnel, celle d'un sommet compact avec de petits contreforts (v. fig. 3, p. 212). Chez les leucémiques, cette figure (v. fig. 2) est plus allongée suivant le grand axe (rapport des axes 3 : 1), et le sommet de la distribution est dévié vers les indices poids-taille élevés. C'est sous ce rapport que se distinguent les distributions des enfants par classes (*Tableau 2*). Dans le groupe d'enfants atteints de leucémie aiguë (garçons et filles) la fréquence des types morphologiquement « moyens » pour les deux caractères est beaucoup plus faible que dans le groupe témoin, alors que la part des enfants ayant des valeurs extrêmes des caractères (M^-) et surtout un rapport perturbé entre le poids et la taille (M^0/M^+, M^+/M^0 et autres) est plus grande.

L'examen de la structure génétique des groupes étudiés a montré qu'on n'observe pas de différences authentiques dans les fréquences géniques des systèmes ABO et Rhésus (*Tableau 3*), bien qu'on puisse noter une certaine augmentation de la fréquence du gène A chez les leucémiques : le rapport des groupes sanguins A : O est de 1,45 chez les malades en comparaison de 1,14 chez les témoins. Des données analogues ont été présentées par d'autres auteurs [5, 8]. L'étude des associations phénotypiques des deux loci révèle des différences certaines dans la structure génétique des enfants comparés : dans le groupe d'enfants sains, la distribution des fréquences des phénotypes correspond théoriquement à celle que l'on espère, en accord avec le

* En ce qui concerne la distribution conjointe du poids et de la taille chez les enfants sains, se référer à la figure 3, p. 212. (*N.d.R.*)

Groupes sanguins des systèmes ABO et Rhésus chez

Système	Groupe sanguin	Enfants atteints de leucémie aiguë						
		N_{obs}	F_{obs}	Fréquence génique	$F_{théor}$	$N_{théor}$	χ^2	H
ABO	O	101	0,2971	$r=0,5547\pm$ $\pm0,0210$	0,3077	104,62	1,39	$0,5822\pm$ $\pm0,0123$
	A	146	0,4294	$p=0,2967\pm$ $\pm0,0193$	0,4172	141,85		
	B	68	0,2000	$q=0,1486\pm$ $\pm0,0143$	0,1869	63,55		
	AB	25	0,0735		0,0882	29,99		
	Total	340						
Rhésus	Rh^+	231	0,8162					$0,4898\pm$ $\pm0,0061$
	Rh^-	52	0,1837	$d=0,4286\pm$ $\pm0,0269$				
	Total	283						

Remarque. N_{obs} est l'effectif observé du phénotype ;
F_{obs} la fréquence observée du phénotype ;
$N_{théor}$ l'effectif théorique du phénotype ;

critère χ^2, alors que dans le groupe d'enfants malades on observe un écart certain de la distribution théorique (*Tableau 4*) et l'augmentation la plus importante de la fréquence du phénotype ARh^+ *.

La distribution des leucémiques étudiés suivant l'âge au début de la maladie se caractérise par une asymétrie très prononcée (fig. 3). La courbe de morbidité est bimodale pour la leucémie. Une « flèche infantile » de la maladie avec maximum à 4 ans et une croissance exponentielle des cas de leucémie sont bien exprimées [9, 10, 12]. Etant donné que nous étudions les enfants jusqu'à 15 ans, nous n'avons affaire qu'à la partie gauche du modèle proposé par Ossetchinski et Fomina [12]. On peut expliquer un certain excès d'enfants de 10 à 15 ans par la présence de la « leucémie des adultes » dans ce groupe d'âge.

* La conclusion de l'augmentation de la fréquence du phénotype ARh^+ chez les leucémiques demande à être vérifiée sur un échantillon plus représentatif statistiquement.

Tableau 3

les enfants atteints de leucémie aiguë et les enfants sains

Enfants sains							Certitude des différences
N_{obs}	F_{obs}	Fréquence génique	$F_{théor}$	$N_{théor}$	χ^2	H	
176	0,3340	$r=0,5785\pm$ $\pm0,0166$	0,3347	176,39	0,01	0,5701± $\pm0,0155$	$\chi^2=2,25$ $P<0,95$
201	0,3814	$p=0,2673\pm$ $\pm0,0148$	0,3807	200,63			
107	0,2030	$q=0,1542\pm$ $\pm0,0115$	0,0824	106,56			
43	0,0816		0,0824	43,42		0,4868± $\pm0,0050$	$\chi^2=0,09$ $P<0,95$
527							
423	0,8246						
90	0,1754	$d=0,4188\pm$ $\pm0,0200$					
513							

$F_{théor}$ la fréquence théorique du phénotype ;
H l'hétérozygotie.

On a également découvert que les caractéristiques génétiques et les indices anthropométriques des enfants à la naissance influencent aussi l'âge dans lequel l'enfant contracte une leucémie aiguë, les enfants ayant des groupes sanguins BRh$^+$ et ARh$^+$ (*Tableau 5*) tombant malades plus jeunes. En ce qui concerne les indices poids-taille, les données obtenues montrent avec évidence que les enfants à poids et taille élevés (M^+) ou bas (M^-) à la naissance contractent plus tôt la maladie et que les enfants à indices poids-taille moyens (M^0) tombent malades plus tard que les autres.

Le *Tableau 6* résume les résultats de la comparaison de certaines caractéristiques des parents pour deux groupes d'enfants. Les parents des leucémiques se sont avérés un peu plus âgés que les parents des enfants sains. Ainsi, seulement 20 % des mères et 31 % des pères de ces derniers avaient plus de 30 ans à la naissance du bébé, tandis que pour ceux des leucémiques ces indices étaient respectivement de 27 et 39 %. Le nombre des grossesses étant égal, celui des accouchements chez les mères des leucémiques est en moyenne plus élevé que chez

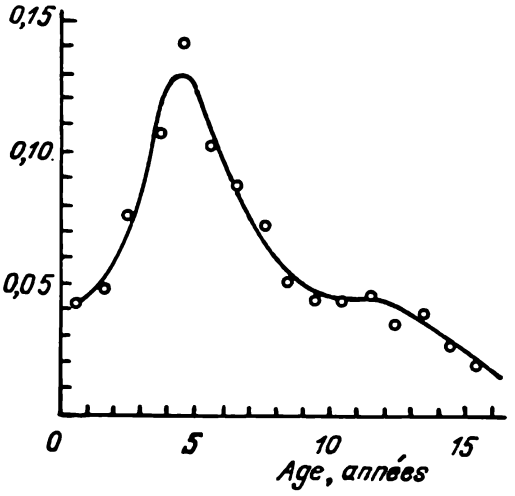


Fig. 3. Distribution en âge (au début de la maladie) des enfants atteints de leucémie aiguë

Tableau 4

Distribution des phénotypes de deux systèmes sanguins chez les leucémiques aigus et les enfants sains

Phénotype	Fréquence théorique du phénotype	Enfants atteints de leucémie aiguë			Enfants sains		
		N_{obs}	F_{obs}	$N_{théor}$	N_{obs}	F_{obs}	$N_{théor}$
ARh^+	0,3139	104	0,3675	88,83	160	0,3119	161,04
ORh^+	0,2760	76	0,2686	78,11	138	0,2690	141,59
BRh^+	0,1667	41	0,1449	47,18	90	0,1754	85,53
$ABRh^+$	0,0679	10	0,0353	19,22	35	0,0682	34,86
ARh^-	0,0668	19	0,0671	18,90	35	0,0682	34,25
ORh^-	0,0587	11	0,0389	16,61	34	0,0663	30,12
BRh^-	0,0355	15	0,0530	10,05	13	0,0253	18,19
$ABRh^-$	0,0145	7	0,0247	4,10	8	0,0156	7,41
N		283			513		
Critère de certitude des différences		$\chi^2 = 14,26$ $P > 0,99$				$\chi^2 = 2,38$ $P < 0,95$	

Remarque. Les désignations sont les mêmes qu'au *Tableau 3*.

Tableau 5

Relation entre l'âge de l'enfant au début de la leucémie aiguë et le groupe sanguin, le poids et la taille à la naissance

Ordre successif de la maladie	Relation entre le système sanguin et l'âge		Relation entre les caractères anthropométriques et l'âge		Critère de certitude des différences
	groupe sanguin	âge, années	classe	âge, années	
I	BRh^+	5,44±0,57	M^-	5,45±0,55	BRh^+-ORh^+
II	ARh^+	6,07±0,36	M^+	5,52±0,48	$P > 0,95$ ($t_d = 2,28$)
	ORh^-	6,00±0,95			
III	ARh^-	6,50±0,93	M^0/M^-	6,14±0,64	M^0-M^-
IV	BRh^-	6,97±1,08	M_{pd}	6,60±0,48	$P > 0,95$ ($t_d = 2,30$)
V	ORh^+	7,15±0,49	M^0	7,10±0,46	M^0-M^-
	$ABRh^+$, $ABRh^-$	7,50±1,00			$P > 0,95$
Moyenne	6,61±0,19 (♂♂ 6,53±0,26, ♀♀ 6,71±0,28)				($t_d = 2,36$)

Remarque. M_{pd} est la classe réunissant les enfants ayant une corrélation perturbée entre le poids et la taille à la naissance (M^0/M^+, M^+/M^0, M^-/M^0, M^+/M^-, M^-/M^+). Les autres désignations sont les mêmes qu'au *Tableau 1* et à la figure 2.

Tableau 6

Age des parents, numéro d'ordre de la grossesse et de l'accouchement dont sont nés les enfants leucémiques et les enfants sains

Certaines caractéristiques des parents	Enfants observés				Critère de certitude des différences
	atteints de leucémie aiguë		sains		
	$M \pm m$	σ	$M \pm m$	σ	
Age de la mère, années	26,61±0,29	5,33	25,47±0,27	5,12	$P > 0,99$ ($t_d = 2,88$)
Age du père, années	28,68±0,32	5,70	28,22±0,64	6,57	$P < 0,95$
Numéro d'ordre de la grossesse	2,15±0,08	1,53	2,17±0,09	1,68	$P < 0,95$
Numéro d'ordre de l'accouchement	1,55±0,05	0,85	1,38±0,03	0,62	$P > 0,99$ ($t_d = 2,91$)
N	323		350		

Remarque. Les désignations sont les mêmes qu'au *Tableau 1*.

les mères des enfants sains. Parmi les enfants sains, les premiers-nés constituent 67 % ; les deuxièmes enfants, 29 % ; les enfants nés d'un troisième accouchement et plus, 4 %. Parmi les enfants devenus par suite leucémiques les premiers-nés représentent 59% ; les deuxièmes enfants, 32 %, les enfants nés d'un troisième accouchement et plus, 9 %.

DISCUSSION

En U.R.S.S., la liaison entre les indices poids-taille des nouveau-nés et la probabilité de contracter par la suite une leucémie aiguë a été étudiée par Vorontzov et coauteurs [1,15]. Selon leurs données, les enfants à indices poids-taille bas ou élevés contractent plus souvent la leucémie. Bernstein et coauteurs [4] ont attiré l'attention sur la tendance à une plus grande fréquence des néoformations malignes chez les enfants nés avec des indices poids-taille élevés. Les résultats que nous avons obtenus permettent, en accord parfait avec les études précédentes [2, 6], de conclure à une résistance maximale à la leucémie aiguë parmi les enfants « moyens » par la totalité des caractères anthropométriques, de même qu'à une prédisposition à cette maladie chez les enfants possédant des valeurs extrêmes des caractères, y compris ceux chez lesquels la corrélation entre le poids et la taille à la naissance est perturbée, c'est-à-dire tous les enfants n'entrant pas dans le domaine de « norme adaptative ». Certes, une telle approche sera encore plus efficace si l'on utilise un nombre plus grand de caractères pour distinguer le domaine de « norme adaptative ».

Lors du choix des caractères, il faut accorder la préférence à ceux qui sont conditionnés pour une grande part par une composante génotypique. On sait que parmi les facteurs déterminant le poids du nouveau-né, le génotype de l'enfant apporte une contribution relative de 18 % de la somme totale des facteurs et le génotype de la mère, de 20 % (Penrose, 1955, cité d'après [14]). L'influence exercée sur les caractères anthropométriques du nouveau-né par des facteurs génotypiques tels que les lots antigéniques de la mère et de l'enfant, la présence d'anomalies chromosomiques chez le fœtus, etc., est bien étudiée. En même temps, on a établi une influence certaine de plusieurs facteurs du milieu (situation sociale de la mère, nombre d'accouchements et d'avortements antérieurs, maladies de la mère, présence de signes pathologiques de la grossesse, etc.). Cependant, leur apport à la détermination du poids à la naissance n'est pas si grand, ce que l'attestent les résultats de l'analyse multifactorielle [7].

Ainsi, bien que le fœtus soit un système « ouvert » et qu'il subisse l'action du milieu extérieur (en premier lieu, du côté de l'organisme maternel), la réponse à une influence de ce genre dépend de la « norme de réaction » du fœtus qui est génétiquement prédéterminée. Ceci signifie que le fœtus est aussi un système « fermé » très autonome

à ses propres lois internes : de nombreuses influences du milieu intervenant comme triggers conditionnent le développement de tel ou tel caractère seulement dans un sens qui est rigoureusement déterminé par l'hérédité. Ce « tamponnement » de l'ontogénèse précoce assure une formation normale, même en présence d'importantes déviations du milieu par rapport à l'optimum [13].

Un des moyens de protection de la formation normale, créés au cours de l'évolution, est la polygénie. Il est évident que derrière les phénotypes ayant des associations optimales de caractères anthropométriques, tout au moins aux stades ontogénétiques précoces, se cachent des complexes déterminés de gènes. L'étude du groupe d'enfants atteints de leucémie aiguë a révélé des différences certaines entre les enfants sains et les leucémiques dans la structure génétique d'après les associations phénotypiques pour deux loci de groupes sanguins ABO et Rhésus. Les différences génétiques augmenteront sans aucun doute si l'on étudie simultanément un grand jeu de loci. Les données relatives aux différences entre les enfants sains et les malades d'après les caractères anthropométriques à la naissance et la structure génétique témoignent de troubles de l'ontogénèse prénatale des individus contractant par la suite une leucémie et, par conséquent, présentant une viabilité relative diminuée. Cette conclusion s'accorde bien avec la fréquence élevée, en comparaison du groupe témoin, des maladies contagieuses et inflammatoires dans les antécédents des enfants devenus leucémiques par la suite (*Tableau 7*). D'autres auteurs [3] ont également signalé un défaut de la protection immunitaire parmi les enfants leucémiques. La confrontation de nos données avec celles de Lozovskaïa et coauteurs [11] révèle une ressemblance frappante entre les fréquences des infections infantiles dans les antécédents des leucémiques et des diabétiques, le diabète sucré étant une maladie à prédisposition héréditaire.

La question de la morbidité par leucémie ne saurait être examinée séparément du problème de la mortalité infantile précoce. La leucémie aiguë est une maladie difficile à guérir et potentiellement fatale. Il a été déjà montré [5] que les enfants morts avant un an, pour différentes causes, se distinguent significativement du groupe témoin par la totalité des caractères anthropométriques et génétiques. A la lumière de ce qui vient d'être exposé, la croissance constatée de la morbidité par leucémie peut s'expliquer par le fait qu'un plus grand nombre d'enfants à viabilité diminuée, qui mouraient autrefois avant un an pour des raisons différentes, atteignent aujourd'hui l'« âge leucémique » et que certains d'entre eux contractent la leucémie. A mesure que la mortalité infantile précoce diminuera, la courbe de début de la leucémie aiguë se déplacera encore davantage vers l'âge plus jeune. Ceci exige l'élaboration de critères diagnostiques permettant le dépistage précoce de la leucémie.

Tableau 7

Maladies contagieuses et inflammatoires subies par les enfants ayant contracté par la suite une leucémie aiguë et un diabète sucré

(sur 100 enfants dans le groupe étudié)

Maladie	Enfants observés					
	Leucémie aiguë, A	Diabète sucré, B *	Sains, C *	*P*		
				A-C	A-B	B-C
Varicelle	48,1	51,7	36,1	$>0,999$	$<0,95$	$>0,999$
Rougeole	37,2	25,0	7,0	$>0,999$	$>0,999$	$>0,999$
Oreillons	30,5	30,0	21,5	$>0,99$	$<0,95$	$>0,999$
Scarlatine	16,7	20,8	11,8	$>0,95$	$<0,95$	$>0,999$
Coqueluche	15,3	12,3	1,6	$>0,999$	$<0,95$	$>0,999$
Rubéole	12,7	10,6	14,0	$<0,95$	$<0,95$	$<0,95$
Hépatite virale	4,0	3,6	0,6	$>0,999$	$<0,95$	$>0,999$
Pneumonie	42,9					
Maladies septicémiques et pyogènes	33,5					
Infections gastro-intestinales	22,1					
N	337	811	811			

* Données de Lozovskaïa et coauteurs [11].

La découverte d'une composante génétique de la prédisposition à la leucémie ne contredit nullement les autres théories de l'étiologie de cette maladie, entre autres la théorie virale. L'existence d'un système génétique coadapté universel suppose sa connexion avec la viabilité relative de l'organisme infantile et avec la résistance non spécifique de ce dernier aux facteurs les plus divers du milieu intérieur et extérieur, les influences leucémigènes entre autres. Il est évident que les données envisagées dans cet article sont un argument de plus en faveur de l'efficacité de l'approche de la génétique des populations, exposée par nous, du problème de la résistance non spécifique de l'organisme humain.

Les auteurs remercient I. Kochel (Institut de pédiatrie de l'Académie des Sciences médicales de l'U.R.S.S., Moscou) pour la possibilité qu'il nous a offerte d'utiliser les archives du service qu'il dirige et pour sa discussion de notre manuscrit.

RÉFÉRENCES

[1] *Алексеев Н. А., Воронцов И. М.* Лейкозы у детей. Л., Медицина, 1979, 174 с.

[2] *Алтухов Ю. П., Курбатова О. Л., Ботвиньев О. К., Афанасьев К. И., Малинина Т. В., Холод О. Н., Стрелкова Л. К., Иванова Н. С.* Генные маркеры и болезни: генетические, антропометрические и клинические особенности детей, больных острой пневмонией. *Генетика*, 1981, 17, 5, 920.

[3] *Бергольц В. М., Кисляк Н. С., Еремеев В. С.* Иммунология и иммунотерапия лейкоза. М., Медицина, 1978, 403 с.

[4] *Берштейн Л. М., Федореев Г. А., Колыгин Б. А., Дильман В. М.* Масса и рост при рождении детей — онкологических больных. *Педиатрия*, 1977, 8, 63.

[5] *Бубнов Ю. И., Кирдан Г. В.* Группы крови системы АВО и их связь с некоторыми злокачественными опухолями у детей. *Генетика*, 1972, 8, 11, 144.

[6] *Дубинин Н. П., Алтухов Ю. П., Курбатова О. Л., Сусков И. И.* Интегральная генетическая характеристика «адаптивной» нормы в популяции человека. *Докл. АН СССР*, 1976, 230, 4, 957.

[7] *Ganguly M., Elwood J. H., McKenzie G.* Factors associated with «low» and «high» birth weight. *Amer. J. Epidemiol.*, 1972, 96, 2, 161.

[8] *Gershowitz H., Neel J. V.* The blood groups and secretor types in five potentially fatal diseases of caucasian children. *Acta genet.*, 1965, 15, 261.

[9] *Кассирский И. А., Алексеев Г. А.* Клиническая гематология. М., Медицина, 1970, 800 с.

[10] *Кисляк Н. С., Махонова Л. А., Ивановская Т. Е.* Клиническое течение и лечение острого лейкоза у детей. М., Медицина, 1972, 208 с.

[11] *Баякина Т. Н.* Значение острых респираторных заболеваний и детских инфекций в развитии сахарного диабета у детей. *Вопр. охраны материнства и детства*, 1980, 25, 7, 20.

[12] *Осечинский И. В., Фомина Р. Ф.* Анализ возрастной зависимости частоты заболеваний острыми лейкозами. *Педиатрия*, 1979, 8, 26.

[13] *Шмальгаузен И. И.* Регуляция формообразования в индивидуальном развитии. М., Наука, 1964, 134 с.

[14] *Штерн К.* Основы генетики человека. М., Медицина, 1965, 204 с.

[15] *Воронцов И. М., Курачева Н. А., Миронович В. К.* Острый лейкоз у детей. Л., Медицина, 1972, 252 с.

NIVEAUX DE POLYMORPHISME ET D'HÉTÉROZYGOTIE DE LA POPULATION RUSSE DE MOSCOU : DONNÉES RELATIVES À 22 LOCI GÉNIQUES CODANT POUR LES PROTÉINES SANGUINES

Y. Altoukhov, R. Khiltchevskaïa et A. Chourkhal

Depuis la publication des premiers travaux [14, 22] ayant permis d'évaluer les niveaux de polymorphisme biochimique et d'hétérozygotie des populations de la drosophile et de l'homme, de nombreuses espèces ont été étudiées par les méthodes de l'électrophorèse des protéines [27]. Les populations de *Homo sapiens*, une des espèces les mieux étudiées sous ce rapport [26], ont été soumises à une analyse particulièrement détaillée. Sur le fond d'un matériel génogéographique aussi vaste, le territoire de l'U.R.S.S. apparaît, en somme, comme une tache blanche, bien que des études adéquates soient la condition essentielle pour comprendre les particularités des phénomènes génétiques dans la population en rapport avec les problèmes de l'urbanisation et de l'état de l'environnement.

Le but de ce travail est de combler cette brèche, ne serait-ce qu'en partie, par l'étude de la population russe de Moscou d'après un ensemble de loci protéiques, suffisant pour obtenir des estimations statistiquement fiables des niveaux de variabilité génétique.

MATÉRIEL ET MÉTHODE

Le matériel a été recueilli en 1976-1979 à l'Institut Central d'Hématologie et de Transfusion sanguine du Ministère de la Santé de l'U.R.S.S. et à la Station de transfusion sanguine de la ville de Moscou. En prenant en considération la composition multinationale de la population moscovite, on n'a tenu compte, en dépouillant les données des formulaires, que des sujets de sexe masculin et de nationalité russe. Le nombre maximal, 1916 donneurs, ont été étudiés pour les loci Tf, Hp, Gc, Hb, LDH, MDH, Est, Alb, 6-PGDH et PGM. 1035 donneurs ont été examinés pour toute la totalité des marqueurs.

Pour l'électrophorèse sur gel de polyacrylamide (GPAA), on utilisait les modifications des systèmes tampons suivant les métho-

© Издательство Наука, 1981.

des [28, 16]: 1° THAM*-EDTA-borate, pH 8,6, THAM 0,115 M, acide borique 0,12 M, EDTA sodique 0,005; 2° THAM-EDTA-borate, pH 8,0, THAM 0,115 M, acide borique 0,16 M, EDTA sodique 0,005 M; 3° THAM-EDTA-borate, pH 8,6, THAM 0,04 M, acide borique 0,04 M, EDTA sodique 0,0015 M.

Pour l'électrophorèse sur gel d'amidon (GA) on utilisait les systèmes tampons suivants (M): 4° acide citrique 0,065, THAM 0,2, EDTA sodique 0,01; 5° histidine-hydrochlorure 0,004, THAM 0,005, EDTA sodique 0,00013; 6° tampon phosphate 0,2.

Pour l'électrophorèse des protéines sur GPAA, on se servait d'appareils mis au point à l'Université de Moscou [4] (épaisseur du gel 4 mm) et des chambres « Havana » et « Desaga » (R.F.A.) (épaisseur du gel 1,5 mm). Pour l'électrophorèse sur GA, on utilisait des chambres verticales de notre propre modification ainsi que la chambre horizontale « Desaga » (R.F.A.).

Les variantes électrophorétiques de protéines étaient identifiées par les techniques suivantes en association avec une coloration spécifique des gels.

I° *Spectre protéique général du sérum sanguin.* GPAA 7,5 %, système tampon d'électrode 2, systèmes tampons de gel 1.

II° *Haptoglobines sériques.* GPAA 6 %, systèmes tampons d'électrode et de gel 1, coloration au 3-amino-9-éthylcarbazole.

III° *Estérase érythrocytaire.* GPAA 7,5 %, système tampon d'électrode et tampon 1.

IV° *Lactate-déshydrogénase.* GPAA 7,5 %, système tampon d'électrode 1, système tampon de gel 3.

V° *Malate-déshydrogénase.* GPAA 7,5 %, système tampon d'électrode 1, système tampon de gel 3.

VI° *Phosphogluconate-déshydrogénase.* GPAA 7,5 %, système tampon 1, système tampon de gel 3, avant l'électrophorèse on ajoute au tampon supérieur 20 mg de NADP **.

VII° *Phosphoglucomutase-1.* GPAA 7,5 % (comme catalyseur de polymérisation on utilisait le diméthylaminopropionitryle), système tampon d'électrode 1, système tampon de gel 3.

VIII° *Phosphatase acide érythrocytaire.* GA 13 %, système tampon d'électrode 4, système tampon de gel 5.

IX° *Adénosine désaminase.* GA 13 %, système tampon d'électrode 4, système tampon de gel 5.

X° *Estérase D.* GA 13 %, système tampon d'électrode 4, système tampon de gel 5.

XI° *Glutamo-pyruvate-transaminase.* GA 13 %, système tampon d'électrode 4, système tampon de gel 5.

XII° *Glyoxalase-1.* GA 10 %, système tampon d'électrode 6,

* THAM ou TRIS: *tris*-(hydroxyméthyl) aminométhane. (*N.d.R.*)
** Nicotinamide-adénine-dinucléotide phosphate. (*N.d.R.*)

système tampon de gel le même à la dilution au 1 : 26e. La coloration des gels pour identifier les différents systèmes enzymatiques, excepté la PGM, a été faite suivant la méthode courante [16] ; pour mettre en évidence l'activité de la PGM sur GPAA, nous avons utilisé notre modification : sel sodique d'acide glucose-1-phosphorique renfermant 1 % de glucose-1,6-diphosphate 2 mg ; glucose-6-phosphate-déshydrogénase 50 U ; NADP 20 mg ; imidazole 2 ou 3 cristaux ; $MgCl_2$ 50 mg ; PMS *1,5 mg ; NBT ** 10 mg ; tampon THAM-HCl, pH 8,0, 50 m.

RESULTATS ET DISCUSSION

La figure montre les électrophorégrammes correspondants de diverses variantes protéiques. Etant donné que, dans certains cas, les techniques que nous avons utilisées se distinguent de celles employées communément pour préparer les solutions tampons et les gels, il convient de signaler les particularités typiques des spectres électrophorétiques (spectres EP) qui se manifestaient alors. Avant tout, ceci concerne les protéines identifiées sur un support de polyacrylamide. Le spectre EP de la LDH dans le système tampon utilisé à pH 8,6 était représenté par 4 fractions principales (v. fig., *a*) : LDH-1, LDH-2, LDH-3, LDH-4. On voit que les principales fractions de l'enzyme sont suivies de sous-fractions mineures à différents degrés d'activité, un spectre particulièrement large étant typique des bandes LDH-2 et LDH-3. En tout, dans le spectre EP de cette enzyme sur GPAA on différencie jusqu'à 12 fractions.

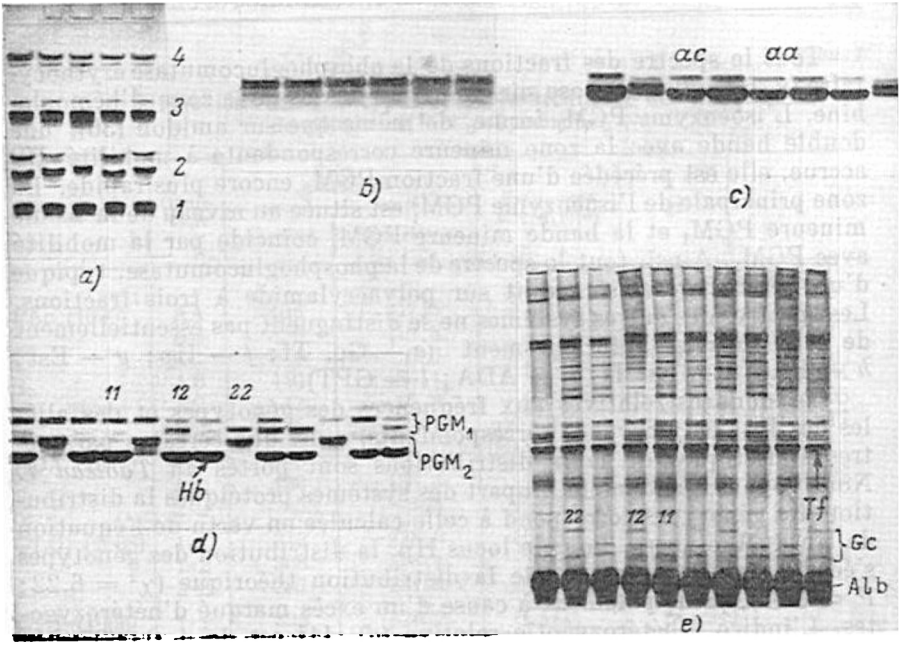
A la différence du type ordinaire de MDH présentant une seule fraction [7], le spectre EP de la malate-déshydrogénase (v. fig., *b*) forme sur GPAA deux bandes dont la zone à grande mobilité possède une activité diminuée. Sur GPAA, l'homozygote (aa) de la 6-PGD (v. fig., *c*) forme aussi deux fractions, ce qui le distingue de la variante mise en évidence par l'électrophorèse sur amidon [3]. La fraction plus mobile possède une activité accrue, les hétérozygotes forment sur GPAA une troisième fraction hybride comme sur un support d'amidon.

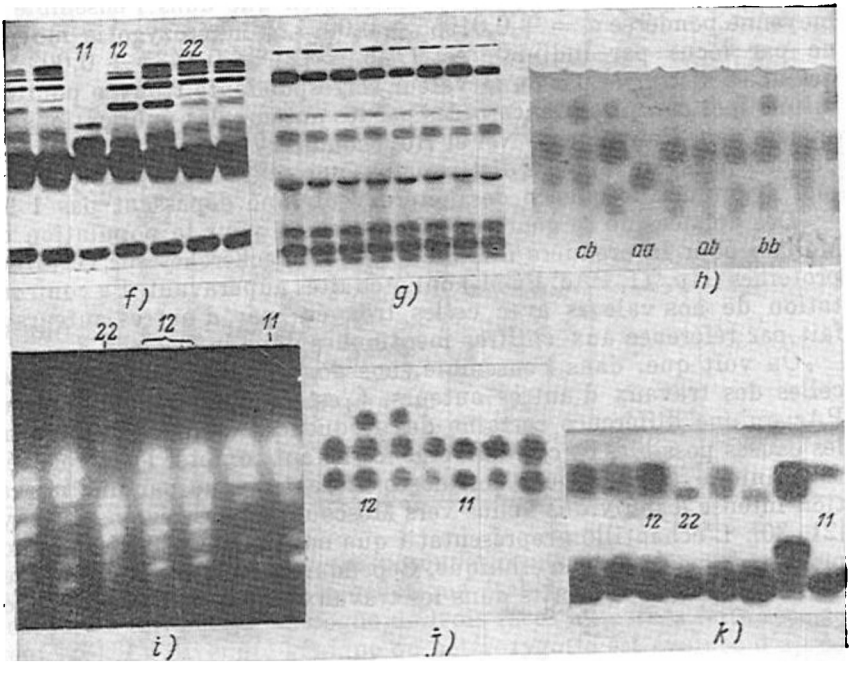
---

* Phénazine méthosulfate. (*N.d.R.*)
** Nitrobleu de tétrazolium. (*N.d.R.*)

Electrophorégramme des protéines érythrocytaires et sériques de l'homme sur gels de polyacrylamide et d'amidon :

Les chiffres et les lettres indiquent les phénotypes (génotypes) : *a* — lactate-déshydrogénase érythrocytaire ; *b* — malate-déshydrogénase érythrocytaire ; *c* — 6-phosphogluconate-déshydrogénase érythrocytaire ; *d* — phosphoglucomutase érythrocytaire ; *e* — protéine totale du sérum ; *f* — haptoglobine sérique ; *g* — estérase ; *h* — phosphatase acide ; *i* — estérase D ; *j* — adénosine-désaminase ; *k* — glutamopyruvate-transaminase


4
3
2
1
a)
b)
αc
αα
c)
11
12
22
PGM1
PGM2
Hb
d)
22
12
11
Tf
Gc
Alb
e)


11
12
22
f)
g)
cb
aa
ab
bb
h)
22
12
11
i)
12
11
j)
12
22
11
k)

Tout le spectre des fractions de la phosphoglucomutase érythocytaire (v. fig., *d*) se dispose sur GPAA au-dessus de la zone d'hémoglobine. L'isoenzyme PGM_1 forme, de même que sur amidon [36], une double bande avec la zone mineure correspondante à mobilité EP accrue, elle est précédée d'une fraction PGM_2 encore plus rapide. La zone principale de l'isoenzyme PGM_1^2 est située au niveau de la bande mineure PGM_1 et la bande mineure PGM_1^2 coïncide par la mobilité avec PGM_2. Ainsi, tout le spectre de la phosphoglucomutase, typique d'un hétérozygote, se réduit sur polyacrylamide à trois fractions. Les spectres des autres systèmes ne se distinguent pas essentiellement de ceux décrits antérieurement (*e* — Gc, Tf; *f* — Hp; *g* — Est; *h* — PAc; *i* — Est D; *k* — ADA; *l* — GPT).

Les données relatives aux fréquences des génotypes et des allèles des loci polymorphes correspondants, ainsi que certains paramètres statistiques de leurs distributions sont portés au *Tableau 1*. Nous voyons que dans la plupart des systèmes protéiques la distribution des génotypes correspond à celle calculée en vertu de l'équation de Hardy-Weinberg. Pour le locus Hp, la distribution des génotypes s'écarte significativement de la distribution théorique ($\chi^2 = 6{,}22$; $P < 0{,}02$), principalement à cause d'un excès marqué d'hétérozygotes. L'indice d'hétérozygotie relative (*d*) [11] est de + 0,0635. Un excès insignifiant d'hétérozygotes est également caractéristique des autres loci, comme PAc, Gc, PGM si bien que dans l'ensemble la moyenne pondérée $d = + 0{,}01955 \pm 0{,}0094$. L'hétérozygotie moyenne par locus par individu est (*Tableau 2*) de $0{,}1435 \pm 0{,}003$, ce qui ne se distingue pas de la valeur correspondante trouvée pour les mêmes loci parmi les Caucasoïdes par les auteurs des publications résumées dans l'ouvrage de Nei et Roychoudhury [26]. La part des loci polymorphes (45 %) ne se distingue pas non plus si l'on y rattache les loci ayant une fréquence des hétérozygotes ne dépassant pas 1 %.

Des données de ce genre ont été obtenues pour la population de Moscou pour la première fois bien que des recherches sur certaines protéines (Hp, Tf, PAc, PGM_1) ont été faites auparavant. La confrontation de nos valeurs avec celles trouvées par d'autres auteurs se fait par référence aux chiffres mentionnés dans le *Tableau 3*.

On voit que, dans l'ensemble, nos données sont en accord avec celles des travaux d'autres auteurs. C'est seulement pour le locus PAc qu'une différence certaine des fréquences est observée. Parmi les causes possibles de cette divergence il faut compter l'hétérogénéité ethnique de la population moscovite entretenue par une migration intense d'individus venus vers Moscou d'autres régions du pays [21, 30]. L'échantillon représentatif que nous avons étudié était homogène du point de vue ethnique. Cependant, il n'est pas sûr que les autres échantillons décrits dans les travaux cités soient aussi homogènes.

Tableau 1

Distribution des génotypes et de la fréquence des allèles de plusieurs loci polymorphes

Locus (*N*)	Génotype	Effectif		Fréquence des allèles	χ^2	*d*
		observé	théorique			
PAc (1030)	AA	106	107,0	0,3223		
	BB	436	440,4			
	CC	1	0,6	0,6539	1,9	
	AB	440	434,1			+0,0134
	AC	12	15,8			−0,2405
	BC	35	32,1	0,0238		+0,0917
ADA (1047)	1-1	903	903,3	0,9288		
	1-2	139	138,4		0,02	+0,0085
	2-2	5	5,3	0,0712		
GLO_1 (1038)	1-1	243	238,0	0,4788		
	1-2	508	518,1		0,39	−0,0194
	2-2	287	281,9	0,5212		
GPT (1006)	1-1	335	329,2	0,5721		
	1-2	481	492,6		0,55	−0,0234
	2-2	190	184,2	0,4279		
Gc (1553)	1-1	642	658,8	0,6513		
	1-2	739	705,4		3,53	+0,0477
	2-2	172	188,4	0,3487		
Hp (1540)	1-1	155	176,9	0,3390		
	1-2	734	690,1		6,22	+0,0638
	2-2	651	673,0	0,6610		
PGM_1 (1559)	1-1	782	784,6	0,7094		
	1-2	648	642,8		0,104	+0,010
	2-2	129	131,6	0,2906		
Est D (965)	1-1	812	811,6	0,9171		
	1-2	146	146,8		0,02	−0,0047
	2-2	7	6,6	0,0829		
PGDH (1553)	a-a	1437	1437,2	0,9620		
	a-c	114	143,5		0,03	+0,0043
	c-c	2	2,3	0,0370		
Tf (1560)	c-c	1516	1515,7	0,9857		
	c-d	20	19,7	0,0065	0,002	+0,015
	c-b	24	23,6	0,0078		+0,013

Les matériaux dont nous disposons permettent également de comparer le niveau de diversité génétique établi pour la population de Moscou à l'estimation analogue de trois races humaines faite récemment [26]. La diversité génique ou hétérozygotie est évaluée d'après

Tableau 2

Hétérozygotie moyenne des populations comparées

Population	Hétérozygotie moyenne pour 22 loci	Références littéraires
Russes (Moscou)	0,1435±0,003	Nos données
Caucasoïdes	0,1419±0,003	[26]
Mongoloïdes	0,1264±0,004	[26]
Négroïdes	0,1149±0,006	[26]

Tableau 3

Fréquence des allèles de certains loci protéiques dans la population russe de Moscou

Allèle	Fréquence de l'allèle		Références littéraires
	Nos propres données	Données d'autres auteurs	
Hp^1	0.339±0,012	0,380±0,025	[35, 29]
PGM_1^2	0,2906±0,0081	0.273±0,061	[19]
PAc^a	0,3223±0,0139	0,3710±0,059	[34]
PAc^b	0,6539±0,0148	0,553±0,063	[34]
Tf^B	0,0156±0,003	0,016±0,007	[42]

les formules $H = \frac{\sum h}{r}$ et $h = (1 - \sum x_i^2)$ où x_i est la fréquence de l'allèle i, r étant le nombre de loci. Pour éviter les divergences, nous ne comparons que les estimations faites pour les mêmes loci (*Tableau 4*).

Il ressort du *Tableau 4* que, pour l'essentiel, la variabilité intéresse les mêmes loci. Cependant, pour les gènes ADA la population russe de Moscou est plus hétérozygote que les trois autres populations comparées. Le locus Hb fait exception, il n'est polymorphe que dans la population négroïde.

Une caractéristique plus exhaustive de la ressemblance et de la différence dans les niveaux de diversité génétique d'après les marqueurs biochimiques étudiés peut être obtenue en utilisant le coefficient de corrélation (r_h) entre les valeurs de l'hétérozygotie à un locus dans les différentes populations (*Tableau 5*).

La plus grande corrélation est constatée dans le couple Russes-Caucasoïdes et la plus petite entre Mongoloïdes et Négroïdes. Par rapport à la population russe, la population mongoloïde occupe une place intermédiaire par la valeur de la corrélation. L'indice considéré

Tableau 4

Distribution des niveaux d'hétérozygotie dans la population moscovite en comparaison de trois races humaines

Locus	Nos propres données	Données d'autres auteurs			Références littéraires
	Russes (Moscou)	Caucasoïdes	Négroïdes	Mongoloïdes	
PAc	0,468±0,0108	0,542±0,013	0,373±0,026	0,322±0,014	[13, 33]
ADA	0,132±0,0137	0,095±0,008	0,058±0,003	0,060±0,006	[18, 32]
Est-D	0,152±0,0148	0,172±0,011	0,164±0,015	0,213±0,004	[4]
GLO	0,499±0,0012	0,493±0,004	0,421±0,017	0,427±0,007	[2]
GPT	0,489±0,0046	0,500±0,001	0,303±0,023	0,481±0,009	[5]
Gc	0,451±0,0075	0,421±0,015	0,159±0,022	0,387±0,009	[1, 39]
Hp	0,448±0,0078	0,484±0,011	0,499±0,003	0,399±0,010	[37, 6]
PGM_1	0,412±0,0096	0,354±0,019	0,320±0,019	0,354±0,006	36, 23]
6-PGDH	0,073±0,0090	0,047±0,006	0,069±0,007	0,125±0,009	[7, 32]
Tf	0,030±0,0031	0,013±0,005	0,085±0,007	0,016±0,004	[37, 12]
Hb	0	0	0,077±0,016	0	[6, 44]

Loci monomorphes

Est_1 [40], Alb [24], Hp_β [38], Hb_α [44, 38], LDH-A et LDH-B [43], MDH [7], PGM_2 [17, 20]

Tableau 5

Coefficients de corrélation entre les valeurs de l'hétérozygotie à un locus des populations comparées

Population	Russes (Moscou)	Caucasoïdes	Mongoloïdes	Négroïdes
Russes (Moscou)				
Caucasoïdes	0,985			
Mongoloïdes	0,950	0,931		
Négroïdes	0,827	0,859	0,794	

reflète une ressemblance qualitative dans la distribution des loci polymorphes parmi les populations comparées. L'évaluation quantitative de la différence génétique par la totalité des loci fournit l'indice de distance (proximité) génétique d'après Nei [25, 26]: $D =$

$= -\ln J_{xy}/\sqrt{J_x J_y}$ où J_{xy} est la moyenne arithmétique entre loci $\sum x_i \cdot y_i$, J_x la moyenne arithmétique $\sum x_i^2$, J_y la moyenne arithmétique $\sum y_i^2$, x_i et y_i étant les fréquences des allèles dans les populations X et Y (*Tableau 6*).

Tableau 6

Distances génétiques entre les populations humaines

Population	Russes (Moscou)	Caucasoïdes	Mongoloïdes	Négroïdes
Russes (Moscou)		0,0031±0,0008	0,0096±0,0030	0,0272±0,0034
Caucasoïdes			0,0116±0,0016	0,0263±0,0040
Mongoloïdes				0,0149±0,0039
Négroïdes				

D'après la grandeur D, le couple Caucasoïdes-Russes est très peu différencié, ce qui concorde avec tous les autres indices génétiques relatifs, caractéristiques de ces populations. La plus grande valeur de la distance (proximité) génétique a été obtenue pour la population moscovite et la population négroïde ($D = 0{,}0272 \pm 0{,}0036$); une valeur intermédiaire de D est obtenue pour le couple Russes-Mongoloïdes. Par la valeur absolue de la proximité génétique la population moscovite est, en comparaison des Caucasoïdes, plus proche des Mongoloïdes et plus éloignée des Négroïdes, mais ces différences ne sont pas certaines.

Tableau 7

Fréquences des variantes électrophorétiques rares de protéines dans la population moscovite en comparaison de données analogues pour les Européens

Locus	Effectif de l'échantillon	Nombre des génotypes rares	Nombre des hétérozygotes rares par 1000 sujets	
			Nos propres données	Données de Harris et coauteurs [15]
PGM_1	1695	3	1,7	1,645
LDH-A	1894	4	2,11	2,956
PGDH	1916	2	1,04	0,400
MDH	1695	2	1,18	3,87
Hb	1695	2	1,18	
Est	1916	2	1,04	

Ainsi, nos données relatives aux niveaux de variabilité génétique d'après les paramètres moyens correspondants dans la population et entre les populations sont en accord avec ce que l'on pourrait espérer *a priori* en tenant compte des particularités de l'ethnogénèse de la population russe.

Nous n'avons non plus mis en évidence aucune particularité des fréquences des variantes électrophorétiques rares de protéines identifiées dans l'échantillon étudié (*Tableau 7*).

Néanmoins, autant que nous sachions, en raison de la pression de migration des gènes venant du dehors, il se déroule d'intenses processus génétiques que l'on peut découvrir en étudiant la structure spatiotemporelle de la population pour les loci des groupes sanguins compte tenu des données démographiques [9, 11, 21, 30]. La tâche de l'avenir consiste à étudier cette question également pour les systèmes de polymorphisme protéique.

RÉFÉRENCES

[1] *Blumberg B. S., Workman P. L., Hirschfeld J.* Gamma-globulin, group specific and lipoprotein groups in a U.S. White and Negro populations. *Nature (Lond.)*, 1964, 202, 561.

[2] *Busi B. R.* Distribution of GLO-1 variants in Western Europe and the Indian subcontinent. *Hum. Genet.*, 1979, 49, 105.

[3] *Carter N. D., Fildes R. A., Fitch L. J., Porr C. W.* Genetically determined electrophoretic variations of human phosphogluconate dehydrogenase. *Acta genet. Basel*, 1968, 18, 109.

[4] *Cartwright R. R., Bethel J. L., Hergreaves H.* The red blood cell esterase D polymorphism in Europe and Asia. *J. Hum. Genet.*, 1976, 33, 161.

[5] *Chen S.-H., Giblett E. R.* Polymorphism of soluble glutamicpyruvic transaminase: a new genetic marker in man. *Science*, 1971, 173, 148.

[6] *Cooper A. J., Blumberg B. S., Workman P. L., McDonough J. R.* Biochemical polymorphic traits in U.S. White and Negro populations. *Amer. J. Hum. Genet.*, 1963, 15, 420.

[7] *Davidson R. G., Cortner J. A.* Genetic variants of human erythrocyte malate dehydrogenase. *Nature (Lond.)*, 1967, 215, 761.

[8] *Davidson R. G.* Electrophoretic variants of human 6-phosphogluconate dehydrogenase: population and family studies and description of a new variant. *Ann. Hum. Genet.*, 1967, 30, 355.

[9] *Дубинин Н. П., Алтухов Ю. П., Курбатова О. Л., Сусков И. И.* Интегральная генетическая характеристика «адаптивной нормы» в популяции человека. *Докл. АН СССР*, 1976, 230, 2, 957.

[10] *Дубинин Н. П., Алтухов Ю. П.* Взаимодействие между преобразованиями окружающей среды и адаптивной демографической и генетической структурой народонаселения. В кн.: Программа «Человек и биосфера» в странах социализма. М., ВИНИТИ, 1979, 195.

[11] *Дубинин Н. П., Алтухов Ю. П.* Мониторинг генетических последствий загрязнения окружающей среды: общий подход. В кн.: Генетические последствия загрязнения окружающей среды. 2. М., Мысль, 1977, 14.

[12] *Giblett E. R.* The plasma transferrins. *Progr. Med. Genet.*, 1962, 2, 34.

[13] *Giblett E. R., Scott M. M.* Red cell acid phosphatase: racial distribution and report of a new phenotype. *Amer. J. Hum. Genet.*, 1965, 17, 425.

[14] *Harris H.* Enzyme polymorphisms in man. *Proc. Roy. Soc.*, 1966, 13, 298.

[15] *Harris H., Hopkinson D. A., Bobson E. B.* The incidence of rare alleles determining electrophoretic variants: data on 43 enzyme loci in man. *Ann. J. Hum. Genet. (Lond.)*, 1974, 37, 237.
[16] *Harris H., Hopkinson D. A.* Handbook of enzyme electrophoresis in human genetics. North-Holland Publ. Co., 1976, 252.
[17] *Hopkinson D. A., Cook P. J. L., Harris H.* Further data on the adenosine desaminase (ADA) polymorphism and a report of a new phenotype. *Ann. Hum. Genet.*, 1969, 32, 361.
[18] *Hopkinson D. A., Harris H.* Rare phosphoglucomutase phenotypes. *Ann. Hum. Genet.*, 1966, 30, 167.
[19] *Ирисова О. В.* Этнографическое распределение типов фосфоглюкомутазы (кф. 2.7.5.1.). Некоторые данные о полиморфизме среди населения Советского Союза. *Вопр. антропол.*, 1977, 56, 45.
[20] *Ishimoto G., Yada S.* Frequency of red cell phosphoglucomutase phenotypes in the Japanese population. *J. Hum. Hered.*, 1969, 19, 198.
[21] *Курбатова О. Л.* Опыт генодемографического исследования больших панмиксных популяций. Генетическая структура двух последовательных поколений жителей Москвы. *Вопр. антропол.*, 1975, 50, 30.
[22] *Lewontin R. C., Hubby J. L.* A molecular approach to the of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics*, 1966, 54, 595.
[23] *Lisker R., Giblett E. R.* Studies on several genetic hematological traits of Mexicans. XI. Red cell acid phosphatase and phosphoglucomutase in three Indian groups. *Amer. J. Hum. Genet.*, 1967, 19, 174.
[24] *Melartin L., Blumberg B. S., Lisker R.* Albumin Mexico, a new variant of serum albumin. *Nature (Lond.)*, 1967, 215, 1288.
[25] *Nei M., Roychoudhury A. K.* Sampling variances of heterozygosity and genetic distance. *Genetics*, 1974, 26, 379.
[26] *Nei M., Roychoudhury A. K.* Genetic variations within and between the major races of man, Caucasoids, Negroids and Mongoloids. *Amer. J. Hum. Genet.*, 1974, 26, 421.
[27] *Nevo E.* Genetic variation in natural populations: patterns and theory. *Theor. Pop. Biol.*, 1978, 13, 121.
[28] *Peackok A. C., Bunting S. C., Queen K. G.* Serum protein electrophoresis in acrilamide gel: patterns from normal human subjects. *Science*, 1965, 147, 1451.
[29] *Филиппов И. К.* Частота типов гаптоглобинов у русского населения г. Москвы. В кн.: Матер. конференций Института медицинской генетики АМН СССР. М., 1972.
[30] *Рычков Ю. Г.* Сравнительное изучение генетического процесса в урбанизированной и изолированной популяциях. *Вопр. антропол.*, 1979, 63, 3.
[31] *Selander H.* Behavior and genetic variation in natural populations. *Amer. Zool.*, 1970, 10, 53.
[32] *Shinoda T.* Polimorphic hematological traits of the Japanes. *Ann. Rep. Natl. Inst. Genet. Japan*, 1969, 20, 101.
[33] *Shinoda T.* A note on the frequency of red cell acid phosphatase types in Japan. *J. Hum. Genet.*, 1979, 13, 249.
[34] *Спицын В. А., Афанасьева И. С., Ирисова О. В., Баева С. Б.* О взаимосвязи между климатическими факторами и географическим распределением генов эритроцитарной кислой фосфатазы человека. *Научн. докл. высшей школы. Биол. науки*, 1977, 3, 68.
[35] *Спицын В. А., Мухина Н. Н.* Распределение типов гаптоглобина у русского населения Архангельской области. *Вопр. антропологии*, 1975, 50, 179.

[36] *Spencer N., Hopkinson D. A., Harris H.* Phosphoglucomutase polimorphism in man. *Nature (Lond.)*, 1964, 214, 742.
[37] *Steinberg A. G., Matsumoto H.* Studies on the Gm, Inv, Hp and Tf serum factors of Japanese populations families. *Hum. Biol.*, 1964, 36, 77.
[38] *Sutton H. E.* Human genetics. *Ann. Rev. Genet.*, 1967, 1, 1.
[39] *Suyama H., Uchida H.* The Gc types in the Japanese population. *Hum. Hered.*, 1969, 19, 74.
[40] *Tashian R. E.* The esterases and carbonic anhydrases of human erythrocytes. In: Biochemical methods in red cell genetics/Ed. by Yunis J.N.Y., Acad. Press, 1969, 307.
[41] *Трувеллер К. А., Нефедов Г. Н.* Многоцелевой прибор для вертикального электрофореза в параллельных пластинах полиакриламидного геля. *Научн. докл. высшей школы. Биол. науки*, 1974, 9, 137.
[42] *Туманов А. Л., Томилин В. В.* Наследственный полиморфизм изоантигенов и изоферментов крови в норме и патологии человека. М., Медицина, 1968, 252.
[43] *Vesel E. S.* Genetic control of isozyme patterns in human tissues. *Progr. Med. Genet.*, 1965, 4, 128.
[44] *Yanase T., Hanada M., Seita M., Ohya I., Ohta Y., Imamura T., Fujimura T., Kawasaki K., Yamaoka K.* Molecular basis of morbidity — from a series of studies of hemoglobinopathies in Western Japan. *Japan. J. Hum. Genet.*, 1968, 13, 40.

LES MARQUEURS GÉNIQUES ET LES MALADIES : PARTICULARITÉS GÉNÉTIQUES, ANTHROPOMÉTRIQUES ET CLINIQUES DES ENFANTS ATTEINTS DE PNEUMONIE AIGUË *

Y. Altoukhov, O. Kourbatova, O. Botviniev, K. Afanassiev, T. Malinina, O. Kholod, L. Strelkova et N. Ivanova

Les connaissances actuelles sur la génétique humaine permettent de conclure que l'une des tâches majeures de la médecine préventive est de comprendre dans quelle mesure la résistance de l'organisme aux maladies est tributaire de ses particularités biologiques individuelles et dans quelle mesure elle est déterminée par des causes extérieures.

Les travaux précédents (v. pp. 208 et 244) ont justifié l'approche de la génétique des populations de la résolution de ce problème et ont montré qu'ayant pour base l'analyse de la variabilité des caractères anthropométriques normaux et des particularités génétiques des nouveau-nés, on peut distinguer les groupes d'enfants plus exposés au danger d'une large gamme de maladies. La présente recherche fournit des preuves complémentaires de l'efficacité de cette approche appuyées sur l'exemple d'une seule maladie: la pneumonie. Cette maladie attire l'attention pour plusieurs causes: haute fréquence des cas et gravité de l'évolution, surtout dans le bas âge; mortalité élevée avant l'usage des sulfamides et des antibiotiques; présence d'une composante génotypique [7].

Le présent travail se distingue des autres publications de ce cycle par un jeu plus large des marqueurs génétiques étudiés. Outre les génotypes des loci polymorphes, on a envisagé les variantes électrophorétiques rares de protéines.

MATÉRIEL ET MÉTHODE

En service hospitalier, on a étudié pendant un an tous les enfants de 6 mois à 4 ans atteints de pneumonie aiguë, 287 en tout (150 ♂♂ et 137 ♀♀). Le diagnostic de la pneumonie était obligatoirement fondé sur les données cliniques, radiologiques et paracliniques. Les enfants malades étaient examinés d'après un programme médico-génétique complexe incluant les points suivants.

* Le travail a été fait avec le secours financier de l'O.M.S., grant G3/181/102.
© Издательство Наука, 1981.

1° Caractéristique clinique [antécédents, poids et taille à la naissance, malformations grandes et petites (stigmates)] ;

2° Analyse sérologique. Détermination des groupes sanguins du système ABO au moyen de sérums standards préparés par l'Institut Central de Transfusion sanguine du Ministère de la Santé de l'U.R.S.S., détermination des groupes sanguins des systèmes MNSs, Rhésus et P au moyen des sérums anti-M, anti-N, anti-S, anti-s, anti-C, anti-C^w, anti-c, anti-D, anti-E, anti-e et anti-P préparés par la firme « Biotest » (R.F.A.).

3° Etude biochimique des protéines érythrocytaires et sériques codées par 11 loci monomorphes et 8 loci polymorphes (v. *Tableau 4* et *5*). Les variantes de protéines ordinaires et rares étaient mises en évidence par électrophorèse sur gels de polyacrylamide (lactate-déshydrogénase, malate-déshydrogénase, phosphogluconate-déshydrogénase et estérase érythrocytaires, hémoglobine, haptoglobine, transferrine, albumine, composante spécifique de groupe de sérum) et d'amidon (phosphoglucomutase, phosphatase acide, superoxyde dismutase et glutamo-pyruvate-transaminase érythrocytaires) suivie de coloration histochimique au Harris-Hopkinson [6] quelque peu modifiée. En cas de découverte de variante protéique rare, on prélevait à l'enfant une deuxième prise de sang et l'on répétait l'analyse biochimique.

Suivant le même programme, 396 nouveau-nés pratiquement sains on été étudiés comme témoins (sans tenir compte des stigmates).

PARTIE EXPÉRIMENTALE

Les données anamnestiques attestent que la plupart des enfants atteints de pneumonie aiguë contractaient souvent, avant leur entrée à l'hôpital, des infections virales respiratoires aiguës (IVRA), des pneumonies (63 %, certains à plusieurs reprises), des otites (16 %) et autres maladies inflammatoires purulentes (24 %). Dans 87 % des cas, une IVRA précédait immédiatement la pneumonie. En service hospitalier, on a mis en évidence chez les enfants observés certaines maladies concomitantes et complications : diathèse exsudative (36 %), anémie hypochrome (22 %), otites (22 %), réactions allergiques aux antibiotiques et autres médicaments (18 %), crises asthmatiformes (10 %), hypotrophie (8 %), infections purulentes (6 %) et autres (7 %) maladies.

L'analyse des malformations et des stigmates montre que 2 enfants souffraient d'une anomalie de la cloison interventriculaire sans troubles circulatoires ; 1 enfant a été opéré dans la période périnatale pour atrésie de l'intestin grêle ; 1 enfant avait le pied bot et l'absence de scrotum. Chez la plupart des enfants, de nombreux stigmates ont été découverts. Leur énumération est donnée au *Tableau 1*.

Tableau 1

Stigmates mis en évidence à l'examen des enfants atteints de pneumonie aiguë

Stigmates	♀♀ N = 93		♂♂ N = 96	
	n	*p*	*n*	*p*
De la tête :				
métopisme	24	0,26	28	0,29
anomalies de l'oreille	6	0,06	18	0,19
palais ogival et semi-ogival	10	0,11	16	0,17
épicanthus	5	0,05	11	0,11
croissance inhabituelle des cheveux sur le front	7	0,07	10	0,10
divers	12	0,13	13	0,13
Des os et des muscles :				
appendice xiphoïde bifide	16	0,17	28	0,29
axiphoïdie	6	0,06	3	0,03
divers	5	0,05	9	0,09
Des téguments :				
atteintes de la région coccygienne	10	0,11	12	0,12
nævus, hémangiomes, etc.	9	0,10	21	0,22
anomalies des lignes principales de la main	14	0,15	5	0,05
pli palmaire transverse fixe et transitoire + + ligne de Sydney	5	0,05	13	0,13
De l'appareil uro-génital :				
cryptorchidie			29	0,30
divers	1	0,01	6	0,06
Total	130	1,40	222	2,32

Remarque. *N* est le nombre d'enfants étudiés ;
n le nombre absolu d'une anomalie morphologique donnée dans le groupe ;
p le nombre d'anomalies par enfant, $p = n/N$.

L'attention est attirée par la fréquence des stigmates, surtout chez les garçons. En moyenne, chaque enfant a 1,5-2,5 stigmates.

Ainsi, le groupe étudié d'enfants atteints de pneumonie aiguë est caractérisé par un niveau de morbidité assez élevé, en particulier de nature virale et bactérienne, ce qui démontre indirectement l'imperfection immunologique de ces enfants. Bien que l'indice unique pour inclure l'enfant dans le groupe étudié soit la pneumonie, la collection de caractères phénotypiques aussi rares que métopisme, anomalies de l'oreille, palais ogival et semi-ogival, cryptorchidie s'est avérée assez grande dans cet échantillon pris au hasard d'enfants atteints de pneumonie. Ceci signifie que la composante traduisant un

trouble d'une certaine résistance non spécifique générale de l'ontogénèse précoce fait un apport de poids à la prédisposition à cette maladie : la pneumonie.

Tableau 2

Paramètres statistiques principaux des distributions des enfants sains et de ceux atteints de pneumonie d'après le poids et la taille à la naissance

Caractère	Paramètres statistiques	Enfants examinés		Critère de certitude des différences
		atteints de pneumonie	sains	
Poids à la naissance, g	$M \pm m$	3260±39	3458±21	$P > 0{,}999$ ($t_d = 4{,}47$)
	σ	663	428	$P > 0{,}999$ ($F = 2{,}40$)
	$\gamma_A \pm m_A$	−0,4838±0,1446	0,3125±0,1231	$P > 0{,}999$ ($t_d = 4{,}19$)
	$\gamma_E \pm m_E$	3,1156±0,2892	3,2185±0,2462	Néant
Taille à la naissance, cm	$M \pm m$	50,10±0,22	51,16±0,09	$P > 0{,}999$ ($t_d = 4{,}46$)
	σ	3,77	1,89	$P > 0{,}999$ ($F = 3{,}98$)
	$\gamma_A \pm m_A$	−1,3896±0,1446	0,2812±0,1228	$P > 0{,}999$ ($t_d = 8{,}81$)
	$\gamma_E \pm m_E$	5,6427±0,2892	2,8829±0,2456	$P > 0{,}999$ ($t_d = 7{,}27$)
	N	287	396	

Remarque. M est la moyenne du caractère, m son erreur, σ l'écart type ; γ_A l'indice d'asymétrie, m_A son erreur ; γ_E l'indice d'excès, m_E son erreur ; t_d le critère de Student ; F le critère de Fisher.

En effet, si nous nous tournons vers les indices anthropométriques — poids et taille à la naissance — des enfants sains et de ceux atteints de pneumonie (*Tableau 2*), nous voyons que la distribution de ces caractères chez les enfants sains suit une loi normale, alors que leur distribution chez les malades présente une asymétrie négative (les deux caractères) et un excès fortement prononcé (la taille). En moyenne, les enfants sains sont plus forts et parmi ceux atteints de pneumonie aiguë la variance des deux caractères est de beaucoup

supérieure. Dans l'ensemble, les différences entre les enfants sains et les malades sont certaines pour 7 paramètres sur les 8 rapportés au *Tableau 2*. Les différences entre les deux groupes d'enfants apparaissent surtout nettement si l'on compare les distributions conjoin-

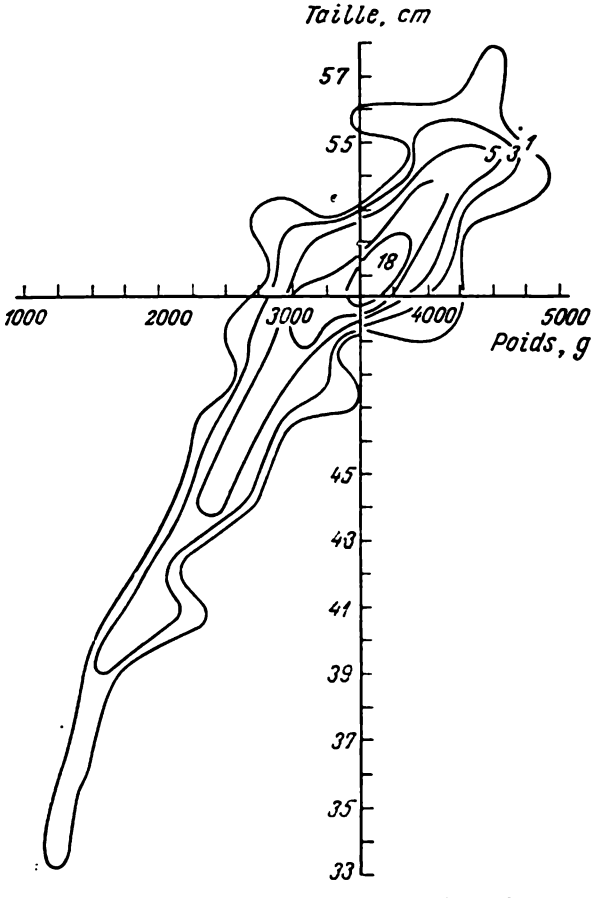


Fig. 1. Distribution conjointe du poids et de la taille à la naissance dans un groupe d'enfants atteints de pneumonie

tes des deux caractères, poids et taille. Le principe de la construction des figures et de la classification des individus suivant deux caractères anthropométriques a été exposé en détail pp. 212-215. La distribution conjointe du poids et de la taille dans le groupe d'enfants sains est représentée sur la figure 3, p. 212. Ce qui caractérise le groupe d'enfants atteints de pneumonie (fig. 1), c'est une forme fortement allon-

gée de la distribution avec prédominance des fréquences dans le domaine aux indices poids-taille bas (M^-) et surtout poids bas (M^0/M^-)*. Le rapport des axes est de 4 : 1. Le *Tableau 3* représente le rapport quantitatif des groupes avec différents indices poids-taille chez les

Tableau 3

Distribution des enfants sains et malades en classes répondant aux caractères anthropométriques

Groupe d'enfants		Classes de caractères anthropométriques										Critère de certitude des différences
		Σ	M^0	M^+	M^-	M^0/M^-	M^0/M^+	M^-/M^0	M^+/M^0	M^+/M^-	M^-/M^+	
Atteints de pneumonie	N	287	51	53	82	49	37	4	7	3	1	$P > 0{,}999$
	%	100	17,8	18,5	28,6	17,1	12,9	1,4	2,4	1,0	0,3	$\chi^2 = 43{,}02$
Sains	N	396	145	85	55	57	38	6	5	4	1	Degrés de liberté = 8
	%	100	36,6	21,5	13,9	14,4	9,6	1,5	1,4	1,0	0,2	

Remarque. Les désignations sont les mêmes que sur la figure 1.

malades et les enfants sains. Dans le groupe d'enfants malades, la part de ceux qui sont morphologiquement « moyens » (pour les deux caractères) est deux fois moindre que dans le groupe témoin, alors que la part des enfants ayant des indices poids-taille bas atteint le maximum : 28,6 % contre 13,9 % parmi les témoins. Ces différences sont d'une certitude élevée. Il est clair que des données de ce genre sont en accord parfait avec la corrélation précédemment démontrée entre les caractères anthropométriques, poids et taille à la naissance, et la résistance à une large gamme de maladies infantiles [1, 4].

On découvre à peu près les mêmes dépendances que celles montrées dans les travaux précédents [4, 5, 11] en comparant la structure génétique des deux groupes d'enfants. En effet, les données présentées au *Tableau 4* montrent que la fréquence des variantes électrophorétiques rares de protéines est à peu près 3 fois plus élevée chez les enfants atteints de pneumonie que chez les enfants sains. Dans le premier cas, on a découvert 7 variantes protéiques rares (fréquence moyenne par locus par individu $9{,}3 \cdot 10^{-4}$). Dans le second, il y en a 4 ($3{,}2 \cdot 10^{-4}$). De plus, il importe d'indiquer que les variantes

* La première désignation se rapporte à la taille ; la seconde, au poids.

Tableau 4

Fréquence des variantes électrophorétiques rares de protéines chez les nouveau-nés sains et les enfants atteints de pneumonie aiguë

Protéines étudiées	Nombre de loci	Enfants examinés			
		atteints de pneumonie aiguë		sains	
		effectif de l'échantillon	nombre d'individus à variante rare	effectif de l'échantillon	nombre d'individus à variante rare
Lactate-déshydrogénase érythrocytaire	2	231	0	364	1 *
Malate-déshydrogénase érythrocytaire	1	231	0	364	0
Phosphogluconate-déshydrogénase érythrocytaire	1	229	4	364	1 *
Estérase érythrocytaire	4	230	1	364	1 *
Phosphoglucomutase érythrocytaire	2	224	1	391	1
Phosphatase acide érythrocytaire	1	175	0	300	0
Superoxyde dismutase érythrocytaire	1	70	0	220	0
Glutamopyruvate-transaminase érythrocytaire	1	95	0	187	0
Hémoglobine	2	231	0	364	0
Transferrine sérique	1	229	0	364	0
Composante spécifique de groupe du sérum	1	230	0	364	0
Albumine sérique	1	230	1	364	0
Fréquence moyenne de la variante rare par locus par individu		0,000926±0,000350		0,000321±0,000161	

* Variantes trouvées chez le même enfant.

de lactate-déshydrogénase, de 6-phosphogluconate-déshydrogénase et d'estérase érythrocytaires identifiées simultanément chez le même nouveau-né du groupe témoin ne sont pas rattachées à un changement de mobilité électrophorétique de la protéine, mais sont plutôt dues à la variabilité régulatrice (fig. 2, *a*, *b*, *c*). En même temps, les variantes découvertes dans le groupe d'enfants malades (v. fig. 2, *d*, *e*, *f*, *g*, *h*, *i*) représentent également des types électrophorétiques pouvant être rangés en catégorie d'hétérozygotes rares. Malheureusement, dans un cas seulement (variante pour la 6-PGDH), on a réussi à examiner simultanément les parents d'un enfant malade. Il s'est avéré que les deux parents avaient un génotype ordinaire pour ce locus (PGDHaa) et que la paternité n'était pas exclue pour 15 loci polymorphes.

Notons que chez les enfants porteurs de variantes protéiques rares, la fréquence des malformations est encore plus élevée qu'en moyenne dans le groupe d'enfants atteints de pneumonie: ainsi chez 7 enfants ayant des variantes protéiques rares, on a diagnostiqué 2 grandes malformations et 30 stigmates (en moyenne, 4,3 par enfant en comparaison de 1,9 dans le groupe d'enfants atteints de pneumonie). En outre, 5 parmi ces 7 enfants avaient à la naissance des indices poids-taille bas. Ceci signifie que les génotypes rares ont une adaptabilité abaissée et, en général, n'entrent pas dans le système reproducteur de la population. Apparemment, la pression de sélection agit le plus fortement contre le groupe « extrême » [3].

En ce qui concerne la distribution des phénotypes ordinaires des loci polymorphes, des différences certaines entre malades et bien portants sont observées pour quatre systèmes génétiques: PGM_1, PAc, MNSs et Rhésus (*Tableau 5*). Pour le locus ABO, bien que les différences n'atteignent pas un niveau de signification, elles sont très probables: chez les enfants malades, la fréquence du gène O croît et celle du gène A décroît. Dans l'ensemble, l'hétérozygotie moyenne par locus dans le groupe d'enfants sains ($0{,}4204 \pm 0{,}0029$) est authentiquement plus élevée que dans celui de malades ($0{,}4058 \pm 0{,}0047$, $P > 0{,}95$). Chez les malades, l'hétérozygotie pour les loci PAc, GPT, ABO est de beaucoup diminuée. Cependant, ce fait ne signifie pas que ce rapport soit typique de chaque locus. Ainsi, pour le locus PGM_1, l'hétérozygotie chez les enfants atteints de pneumonie aiguë est plus élevée de façon certaine que chez les enfants sains ($P > 0{,}999$). D'après les groupes sanguins, les malades atteints de pneumonie se distinguent, en outre, des enfants sains par une fréquence élevée des associations antigéniques rares. Dans le système MNSs, la fréquence du chromosome NS est élevée, tandis que dans le système Rhésus, la fréquence globale des chromosomes rares ($< 10\%$) (CDE, CdE, Cde, cdE, C^wDE, C^wDe, C^wde) chez les enfants sains est de 0,08 et chez les enfants atteints de pneumonie aiguë, de 0,13 ($P > 0{,}95$).

Caractéristique des enfants sains et atteints de

Locus	Génotype	Enfants atteints de pneumonie					
		N_{obs}	$N_{théor}$	Fréquence de l'allèle	χ^2	H	D
6PGDH	a	168	167,16	$PGDH^a = 0,9718$	0,08	$h_{obs} = 0,0452$	−0,1752
	ac	8	9,70	$PGDH^c = 0,0282$		$H_{théor} = 0.0548 \pm$	
	c	1	0,14			$\pm 0,0166$	
	N	177					
PGM_1	1-1	64	69,30	$PGM^1 = 0,6329$	3,02	$h_{obs} = 0,5260$	0,1319
	1-2	91	80,39	$PGM^2 = 0,3671$		$H_{théor} = 0.4647 \pm$	
	2-2	18	23,32			$\pm 0,0139$	
	N	173					
SOD A	1-1	46	46,01	$SOD^1 = 0,9894$	0	$h_{obs} = 0,0213$	0,0143
	1-2	1	0,99	$SOD^2 = 0,0106$		$H_{théor} = 0.0210 \pm$	
	2-2	0	0			$\pm 0,0205$	
	N	47					
PAc	A	7	7,17	$PAc^a = 0,2313$	2,41	$h_{obs} = 0,3956$	−0,0232
	B	74	72,41	$PAc^b = 0,7351$		$H_{théor} = 0.4050 \pm$	
	C	0	0,15	$PAc^c = 0,0336$		$\pm 0,0288$	
	AB	44	45,57				
	BC	5	6,62				
	AC	4	2,08				
	N	134					
GPT	1-1	17	16,04	$GPT^1 = 0,5259$	0,26	$h_{obs} = 0,4655$	−0,0666
	1-2	27	28,92	$GPT^2 = 0,4741$		$H_{théor} = 0.4987 \pm$	
	2-2	14	13,04			$\pm 0,0077$	
	N	58					
Tf	c	174	174,01	$Tf^c = 0,9915$	0	$h_{obs} = 0,0169$	0
	cb	1	0,99	$Tf^b = 0,0028$		$H_{théor} = 0.0169 \pm$	
	cd	2	2	$Tf^d = 0,0057$		$\pm 0,0096$	
	N	177					
Hp	1-1	28	28,73	$Hp^1 = 0,4017$	0,05	$h_{obs} = 0,4888$	0,0169
	1-2	87	85,56	$Hp^2 = 0,5983$		$H_{théor} = 0.4807 \pm$	
	2-2	63	63,72			$\pm 0,0104$	
	N	178					
ABO	O	81	83,16	$p = 0,2265$	1,54	$H_{théor} = 0.5390 \pm$	
	A	74	71,59	$q = 0,1516$		$\pm 0,0200$	
	B	48	45,47				
	AB	12	14,77	$r = 0,6219$			
	N	215					

Tableau 5

pneumonie d'après la totalité des marqueurs génétiques

Enfants sains						Différences entre les enfants sains et malades en unités χ^2
N_{obs}	$N_{théor}$	Fréquence de l'allèle	χ^2	H	D	
336 13 1 350	335,19 14,66 0,17	$PGDH^a = 0,9786$ $PGDH^c = 0,0214$	0,05	$h_{obs} = 0,0419$ $H_{théor} = 0,0419 \pm$ $\pm 0,0105$	0	0,45
188 173 30 391	192,68 163,59 34,72	$PGM^1 = 0,7020$ $PGM^2 = 0,2980$	1,30	$h_{obs} = 0,4425$ $H_{théor} = 0,4184 \pm$ $\pm 0,0132$	0,0576	6,14 *
217 3 0 220	214,01 2,97 0,02	$SOD^1 = 0,9932$ $SOD^2 = 0,0068$	0	$h_{obs} = 0,0132$ $H_{théor} = 0,0135 \pm$ $\pm 0,0077$	0	0,15
35 97 1 124 31 12 300	35,37 101,52 1,68 119,82 26,19 15,45	$PAc^a = 0,3433$ $PAc^b = 0,5817$ $PAc^c = 0,0750$	2,28	$h_{obs} = 0,5566$ $H_{théor} = 0,5381 \pm$ $\pm 0,0128$	0,0344	23,59 ***
47 107 33 187	54,00 92,98 40,02	$GPT^1 = 0,5374$ $GPT^2 = 0,4626$	4,25 *	$h_{obs} = 0,5722$ $H_{théor} = 0,4972 \pm$ $\pm 0,0043$	0,1580	2,19
346 3 5 354	346,03 2,94 4,96	$Tf^c = 0,9387$ $Tf^b = 0,0042$ $Tf^d = 0,0071$	0	$h_{obs} = 0,0226$ $H_{théor} = 0,0224 \pm$ $\pm 0,0078$	0	0,20
60 222 183 465	62,87 216,22 185,91	$Hp^1 = 0,3677$ $Hp^2 = 0,6323$	0,33	$h_{obs} = 0,4774$ $H_{théor} = 0,4650 \pm$ $\pm 0,0084$	0,0267	1,31
113 139 73 30 335	113,69 138,23 72,19 30,90	$p = 0,2765$ $l = 0,1577$ $r = 0,5659$	0,04	$H_{théor} = 0,5784 \pm$ $\pm 0,0127$		3,84

Locus	Génotype	Enfants atteints de pneumonie					
		N_{obs}	$N_{théor}$	Fréquence de l'allèle	χ^2	H	D
MNSs	MS	8	9,16	MS = 0,2334	13,78*	$h_{obs} = 0,6846$	−0,0451
	MSs	28	24,64	Ms = 0,3142		$H_{théor} = 0,7169 \pm$	
	Ms	15	16,58	NS = 0,1088		$\pm 0,0082$	
	MNS	12	8,54	Ns = 0,3436			
	MNSs	32	38,43				
	MNs	38	36,27				
	NS	5	1,99				
	NSs	5	12,56				
	Ns	25	19,83				
	N	168					
Rhésus	CCDEE	2	0,13	CDE = 0,0238	17,70	$H_{théor} = 0,7656 \pm$	
	CCDEe	1	4,23	CDe = 0,3830		$\pm 0,0133$	
	CcDEE	0	1,55	cDE = 0,1473			
	CCDee	33	35,35	cDe = 0,1469			
	CcDEe	37	29,02	CdE = 0,0030			
	CcDee	53	50,48	Cde = 0,0678			
	ccDEE	3	4,75	cdE = 0,0169			
	ccDEe	17	18,94	cde = 0,1973			
	ccDee	14	14,16	C^wDE = 0,0028			
	CCddEE	0	0	C^wDe = 0,0056			
	CCddEe	0	0,07	C^wde = 0,0056			
	CcddEE	0	0,02				
	CcddEe	1	0,62				
	CCddee	0	0,82				
	Ccddee	5	4,76				
	ccddEE	0	0,05				
	ccddEe	1	1,18				
	ccddee	6	6,93				
	C^wCDee	2	1,66				
	C^wcDEE	1	0,16				
Rhésus	C^wcDEe	0	0,96				
	C^wcDee	0	0,98				
	$C^wccddee$	2	0,39				
	N	178					
P	P+	87		$p = 0,5514$		$H_{théor} = 0,4947 \pm$	
	P−	38				$\pm 0,0070$	
	N	125					

Remarque. 1° N_{obs} est l'effectif observé des phénotypes, $N_{théor}$ l'effectif théorique observée, D l'hétérozygotie relative $D = \frac{h_{obs} - H_{théor}}{H_{théor}}$.

2° Les astérisques indiquent les différences certaines au niveau de signification

3° Etant donné que chez les nouveau-nés la maturation ontogénétique du système Hp

4° Pour évaluer les fréquences des gènes MNSs et Rhésus, nous avons utilisé la méthode gage Fortran (programme de traduction « Doubna ») par B. Sémënov, collaborateur du labo- Sciences de l'U.R.S.S. avec déchiffrement du programme aimablement présenté par le pro-

Suite

Enfants sains						Différences entre les enfants sains et malades en unités χ^2
N_{obs}	$N_{théor}$	Fréquence de l'allèle	χ^2	H	D	
20	18,27	MS = 0,2888	11,29 *	$h_{obs} = 0,6941$	−0,0240	15,79 *
36	44,62	Ms = 0,3527		$H_{théor} = 0,7112 \pm$		
30	27,25	NS = 0,0879		$\pm 0,0069$		
19	11,12	Ns = 0,2705				
44	47,80					
46	41,80					
0	1,69					
7	10,42					
17	16,03					
219						
0	0,01	CDE = 0,0023	10,96	$H_{théor} = 0,7452 \pm$		32,70 *
2	1,51	CDe = 0,3094		$\pm 0,0100$		
0	0,87	cDE = 0,1807				
27	26,74	cDe = 0,0906				
35	29,58	CdE = 0,0086				
57	60,11	Cde = 0,0426				
5	7,70					
33	34,39					
16	15,30					
0	0,02					
0	0,16					
0	0,03					
2	1,42	cdE = 0,0070				
1	0,40	cde = 0,3405				
5	6,35	$C^wDe = 0,0183$				
0	0,01					
1	1,05					
27	25,39					
0	2,82					
0	0					
4	1,50					
4	3,45					
0	0					
219						
186		$p = 0,5932$		$H_{théor} = 0,4826 \pm$		0,89
101				$\pm 0,0077$		
287						

des phénotypes, H l'hétérozygotie, $H_{théor}$ l'hétérozygotie théorique, h_{obs} l'hétérozygotie

* $P > 0,95$, ** $P > 0,99$, *** $P > 0,999$.

n'a pas encore eu lieu, les données correspondantes pour leurs mères sont mentionnées.

de la vraisemblance maximale [12]. Le programme a été adapté pour ЕЭСМ-6 dans le laboratoire de génétique des populations de l'Institut de Génétique générale de l'Académie des fesseur J. Neel (chaire de génétique humaine de l'Université du Michigan, U.S.A.).

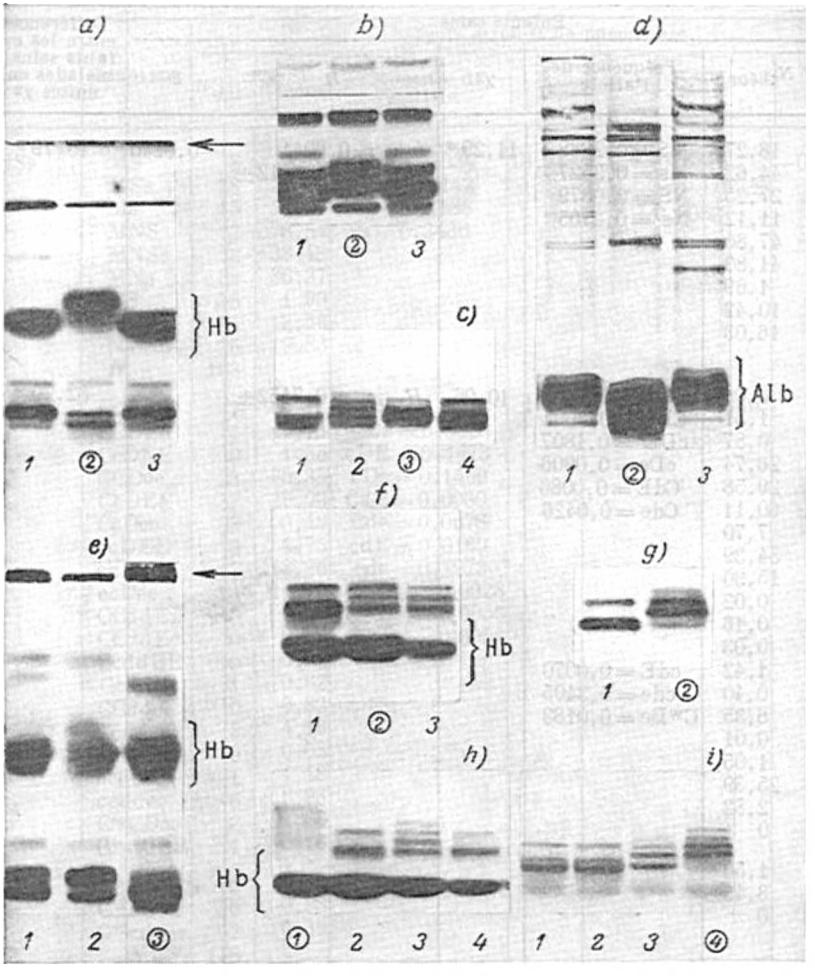


Fig. 2. Exemples de variantes électrophorétiques rares de protéines (chiffres encerclés) découvertes lors de l'examen des prises de sang d'enfants sains (*a-c*) et d'enfants atteints de pneumonie (*d-i*) :

a et *e* — estérase érythrocytaire ; *b* — lactate-déshydrogénase érythrocytaire ; *d* — albumine sérique ; *c*, *f*, *g*, *h*, *i* — phosphogluconate-déshydrogénase érythrocytaire ; *c* : n° 1 — 4 homozygotes PGDH-aa ; n° 2 — hétérozygote ordinaire ac ; n° 3 — variante rare ; *f* : n° 1 — homozygote aa ; n° 3 — hétérozygote ac ; n° 2 — variante rare ; *g* : n° 1 — homozygote aa ; n° 2 — variante rare ; *h* : n° 1 — variante rare ; n°s 2 et 4 — homozygotes aa ; n° 3 — hétérozygote ac ; *i* : n°s 1 et 2 — homozygotes aa ; n° 3 — hétérozygote ac ; n° 4 — variante rare. Sur tous les clichés, l'anode est en bas, la cathode, en haut. Electrophorèse sur gel de polyacrylamide, système tampon THAM-EDTA, pH 8,6. Alb — albumine, Hb — hémoglobine. La flèche indique le start

DISCUSSION

Malgré toute l'évidence de l'adage « il faut soigner non pas la maladie, mais le malade », jusqu'à présent on ne prête pas assez d'attention à cette circonstance que chaque malade fait partie d'une population possédant un patrimoine génétique spécifique. Même si l'on ne tient compte du type de processus génétique — normal ou défavorable — qui se déroule dans la population étudiée, on voit clairement que chaque génération refuse les génotypes les moins adaptés biologiquement. Peu importe, de plus, ce qui provoque l'apparition de génotypes aussi rares : une mutation *de novo* [5] ou une ségrégation et une recombinaison des allèles de loci polymorphes d'après certains complexes coadaptés de gènes [4, 8]. Dans l'un et l'autre cas, les conséquences biologiques peuvent être les mêmes.

Comme nous l'avons déjà montré [4], les génotypes moins bien adaptés se concentrent généralement aux limites des distributions des caractères anthropométriques et il s'établit ainsi un lien entre la génétique des caractères quantitatifs et celle des traits mendéliens ainsi qu'entre ces particularités et la morbidité. Sans parler de nombreux autres loci des groupes sanguins pour lesquels la pression de sélection est prouvée de façon convaincante [2, 10, 12], même des marqueurs génétiques semblerait-il neutres comme certaines variantes électrophorétiques de protéines peuvent être impliqués dans une sélection pour la résistance non spécifique de l'organisme aux maladies. Donc, les efforts visant à supprimer les causes extérieures de la maladie peuvent être peu efficaces ou même tout à fait inopérants si l'on ne sait rien du génotype. C'est la raison pour laquelle l'approche de la génétique des populations doit indiscutablement faire partie de n'importe quel complexe de mesures de la médecine préventive.

Les auteurs remercient le professeur J. Neel (chaire de génétique humaine de l'Université du Michigan, U.S.A.) du programme offert pour les calculs des fréquences génétiques des systèmes polymorphes, et B. Séménov (laboratoire de génétique des populations de l'Institut de Génétique générale de l'Académie des Sciences de l'U.R.S.S.) pour l'exécution des calculs.

RÉFÉRENCES

[1] *Алтухов Ю. П.* Генетический мониторинг популяций в связи с состоянием среды. XIV Международный генетический конгресс. Пленарные заседания, симпозиумы. Тезисы докладов, М., Наука, 1978, 16.

[2] *Cavalli-Sforza L. L., Bodmer W. F.* The genetics of human populations. San Francisco, W. H. Freeman and Co., 1971, 252.

[3] *Dubinin N. P., Altukhov Yu. P.* Gene mutations (*de novo*) found in electrophoretic studies of blood proteins of infants with anomalous development. *Proc. Nat. Acad. Sci. U.S.A.*, 1979, 76, 10, 5226.

[4] *Дубинин Н. П., Алтухов Ю. П., Курбатова О. Л., Сусков И. И.* Интегральная генетическая характеристика « адаптивной нормы » в популяциях человека. *Докл. АН СССР*, 1976, 230, 4, 957.

[5] *Дубинин Н. П., Алтухов Ю. П., Сусков И. И., Хильчевская Р. И., Афанасьев К. И., Малинина Т. В., Большакова Н. К., Шурхал А. В., Курбатова О. Л., Прохоровская В. Д.* Экспериментальное обоснование принципов мониторинга генных мутаций у человека. *Докл. АН СССР*, 1978, 243, 5, 1313.

[6] *Harris H., Hopkinson D. A.* Handbook of enzyme electrophoresis in human genetics. North-Holland Publ. Co., Amsterdam, Oxford, N.Y., American Elsevier Publ. Co. Inc., 1976.

[7] *Канаев И. И.* Близнецы и генетика. Л., Наука, 1968, 76.

[8] *Курбатова О. Л.* Опыт генодемографического исследования больших панмиксных популяций. Генетическая структура двух последовательных поколений жителей Москвы. *Вопр. антропологии*, 1975, 50, 30.

[9] *Reed T. E., Schull W. J.* A general maximum likelihood estimation program. *Amer. J. Hum. Genet.*, 1968, 20, 6, 579.

[10] *Рычков Ю. Г., Русакова О. Л., Раппопорт М. П., Шереметьева В. А.* Факторы генетической дифференциации популяционной системы коренного населения Северной Азии. Сообщение I. *Генетика*, 1973, 9, 2, 136.

[11] *Сусков И. И., Алтухов Ю. П., Афанасьев К. И., Курбатова О. Л., Малинина Т. В.* Редкие варианты мономорфных и полиморфных белковых локусов у детей в норме и патологии. IV Всесоюзный симпозиум «Молекулярные механизмы генетических процессов». Тезисы докладов. М., 1979, 75.

[12] *Yamazaki T., Maruyama T.* Evidence that enzyme polymorphisms are selectively neutral, but blood groupe polymorphisms are not. *Science*, 1974, 183, 1091.

LA MIGRATION COMME FACTEUR D'HÉTÉROGÉNÉITÉ GÉNÉTIQUE DE LA POPULATION DES GRANDES VILLES

B. Kalabouchkine, V. Prokhorovskaïa, R. Khiltchevskaïa et Y. Altoukhov

Sous l'aspect génétique, la population d'une grande ville ne s'autoreproduit pas, car le maintien et surtout la croissance de l'effectif se font principalement aux dépens de l'afflux de migrants [8, 10]. Ce processus s'accompagne de la destruction de la structure de populations historiquement formée et entraîne une augmentation de la recombinaison génique [1, 2, 5, 6]. Des phénomènes de ce genre dans les populations humaines ne sont pratiquement pas étudiés au point de vue génétique, bien que l'importance de leurs données, pour le problème de l'adaptation de la population à un environnement auquel elle n'a pas eu affaire dans son passé historique, soit difficile à surestimer.

Le but de cette recherche — partie intégrante du programme des travaux de l'Institut de Génétique générale de l'Académie des Sciences de l'U.R.S.S. dans le cadre du problème « L'Homme et la Biosphère » — est d'étudier dans quelle mesure le facteur migration influe sur l'hétérogénéité génétique de la population de Moscou. Ces données sont importantes non seulement par elles-mêmes, mais aussi en rapport avec la justification des différentes approches de la solution des problèmes du monitoring génétique des populations humaines dans les conditions de la pollution et d'autres variations de l'environnement.

MATÉRIEL ET MÉTHODE

Le travail a été effectué en 1976-1979. Le schéma de l'analyse comprenait la détermination des groupes sanguins du système ABO chez les donneurs venus pour la première fois et examinés à l'Institut de Transfusion sanguine du Ministère de la Santé publique de l'U.R.S.S., ainsi que le choix de certaines données démographiques concernant les donneurs eux-mêmes et leurs parents. On a utilisé dans ce travail les résultats du recensement de la population permettant d'évaluer

© Издательство Наука, 1982.

l'afflux de migrants à Moscou au cours des périodes biennales précédentes [7]. En outre, nous avons utilisé les données sur la distribution des fréquences des groupes sanguins du système ABO dans la population moscovite, englobant à peu près une période de cinquante ans. Nos matériaux [9] ainsi que les données de Grigorova [cité d'après 11] et d'Oumnova et collaborateurs [12] ne comprennent la détermination des groupes sanguins parmi les donneurs que de la population russe de Moscou. Dans les travaux de Rosanova [cité d'après 11], Bachlaï [3] et Kourbatova [8] les groupes sanguins étaient déterminés parmi la population de composition nationale mixte de Moscou.

Afin d'évaluer la similitude génétique des distributions des groupes sanguins présentées par des auteurs différents, nous avons utilisé l'indice de similitude r et le critère d'identité I [13]:

$$r = \sqrt{p_1 q_1} + \sqrt{p_2 q_2} + \ldots + \sqrt{p_m q_m}$$

où $p_1, p_2, \ldots, p_m$ sont les fréquences des phénotypes dans le premier échantillon,

$q_1, q_2, \ldots, q_m$ les fréquences des phénotypes dans le deuxième échantillon.

Le critère d'identité I est utilisé pour estimer la valeur de r:

$$I = \frac{8 N_1 N_2}{N_1 + N_2} (1 - r)$$

où N_1 et N_2 sont les effectifs correspondants dans le premier et le deuxième échantillon,

I est distribué comme χ^2 avec $m - 1$ degrés de liberté.

Pour comparer plusieurs échantillons, on a utilisé l'indice de similitude généralisé R:

$$R = \sqrt[k]{p_{11} p_{12} p_{1k}} + \ldots + \sqrt[k]{p_{m1} p_{m2}, \ldots, p_{mk}}$$

où k est le nombre d'échantillons comparés,

m le nombre de phénotypes,

p_{iJ} la fréquence du phénotype i dans l'échantillon J.

Afin d'estimer la valeur statistique, on a calculé l'erreur d'échantillonnage approchée:

$$m_R = \sqrt{\frac{1 - R^2}{k \tilde{N}}}$$

où

$$\frac{1}{\tilde{N}} = \frac{1}{k} \sum_{1}^{k} \frac{1}{N_J}.$$

PARTIE EXPÉRIMENTALE ET DISCUSSION

D'après le recensement de 1970, l'afflux de migrants à Moscou au cours de la période biennale précédente représentait 2,9 % de tout l'effectif de la population y habitant en permanence. En se basant sur ces matériaux et sur d'autres, on peut s'attendre à ce que la population de Moscou au cours d'une génération ($\approx$ 30 ans) augmente de 45 % en moyenne aux dépens de personnes venues des différentes régions du pays. Certains travaux renferment des données relatives à la structure démographique de la population de Moscou. Par exemple, d'après les estimations de Kourbatova [8] parmi les personnes nées en 1936-1966 le pourcentage des mariages conclus entre Moscovites est de 30,9 % ; le pourcentage des couples où seul un conjoint est Moscovite est de 31,9 % ; la fréquence des cas où les deux conjoints sont des migrants est de 37,2 %. Le *Tableau 1* présente les données de la structure matrimoniale des couples de parents des donneurs examinés par nous.

Tableau 1

Structure matrimoniale des couples de parents des donneurs examinés (en %)

Année de la recherche	Années de naissance des parents des donneurs	Les deux conjoints étaient nés à Moscou ou dans les environs	Un seul conjoint était né à Moscou ou dans les environs	Les deux conjoints étaient des migrants
1976	1894 à 1934	25,1	24,3	50,6
1978-1979	1902 à 1937	21,4	15,1	63,5

Ainsi, des sources indépendantes révèlent que la part des couples venus à Moscou d'autres régions du pays peut aller de 37 à 63 % dans des générations différentes. Comme la migration provient de régions différant par les paramètres génétiques de la population [4] et n'est pas limitée par l'appartenance nationale, elle peut être un facteur de l'hétérogénéité génétique de la population de Moscou. En effet, Kourbatova [8] et nous-mêmes plus tard [9] avons constaté une telle hétérogénéité.

On peut arriver à la même conclusion en analysant les publications englobant près d'un demi-siècle d'étude des groupes sanguins du système ABO de la population de Moscou. Le *Tableau 2* montre les fréquences des phénotypes et des gènes calculées avec la correction de Bernstein sur la base des données de plusieurs auteurs. Les distributions des fréquences des phénotypes présentées au *Tableau 2* diffèrent de façon certaine les unes des autres ($R = 0{,}9970 \pm 0{,}0012$, $P > 0{,}004$).

Tableau 2

Fréquence des phénotypes et des gènes du locus ABO dans la population de Moscou

O	A	B	AB	r	p	q	N	Références littéraires
0,333	0,374	0,228	0,065	0,5872	0,2526	0,1602	2 740	1931 [9]
0,337	0,375	0,209	0,079	0,5822	0,2614	0,1564	10 116	1947 [9]
0,335	0,378	0,206	0,081	0,5791	0,2653	0,1556	31 896	1968 [10]
0,340	0,377	0,206	0,077	0,5837	0,2624	0,1539	63 549	1968 [11]
0,343	0,369	0,209	0,079	0,5868	0,2588	0,1544	44 080	1975 [1]
0,342	0,385	0,184	0,089	0,5783	0,2740	0,1476	890	Données des auteurs pour 1976-1979

Le test traditionnel portant sur l'hétérogénéité — critère χ^2 — révèle le même degré d'hétérogénéité ($\chi^2 = 28{,}4$, $P < 0{,}005$). Remarquons que l'estimation de l'indice de similitude s'avère très proche des valeurs de R calculées pour la population russe de la partie européenne de la R.S.F.S.R. ($R = 0{,}9952 \pm 0{,}00065$) et pour la population autochtone de la République socialiste soviétique d'Ukraine ($R = 0{,}9981 \pm 0{,}00013$). Le *Tableau 3* donne les valeurs des critères de similitude (r) et d'identité (I) pour des grandeurs du *Tableau 2* comparées deux par deux. On voit que les estimations de la similitude entre les données présentées par différents auteurs ne dépendent pas fortement de l'intervalle de temps, mais seraient plutôt d'un caractère fortuit, ce qui peut mener à conclure à la présence d'une hétérogénéité génétique de la population.

Etant donné que les recherches de différents auteurs ont été effectuées sur des groupes de population dissemblables pour ce qui est de la composition nationale, on pourrait croire que l'hétérogénéité observée est précisément liée à cette circonstance.

Nous avons comparé les données pour les groupes de composition nationale mixte et pour la population russe seule. L'indice de similitude entre les groupes de composition nationale mixte: $R_1 = 0{,}99997 \pm 0{,}000030$, c'est-à-dire qu'il ne se distingue pas de 1 avec certitude. L'indice de similitude entre les groupes d'origine russe ($R_2 = 0{,}99881 \pm 0{,}000534$) révèle des différences certaines entre eux. En outre, les valeurs de R_1 et de R_2 obtenues diffèrent entre elles de façon certaine: $P < 0{,}001$. L'estimation des différences a été faite après les conversions correspondantes recommandées pour les valeurs élevées de R [13]. Comme il fallait s'y attendre, ce résultat est en accord parfait avec les données du *Tableau 3*. Les valeurs de

l'indice de similitude s'avèrent plus basses quand on compare les groupes de population russe, alors qu'elles sont plus élevées en comparant les groupes de composition nationale mixte. Nous ne discuterons pas les causes possibles des différences constatées entre les valeurs de R_1 et R_2. Cependant, il est évident que l'hétérogénéité

Tableau 3

Valeurs des critères de similitude (r) et d'identité (I) * dans la comparaison des distributions des fréquences des groupes sanguins du système ABO dans la population de Moscou

(d'après les données de divers auteurs, v. *Tableau 2*)

Années de la recherche	1931	1947	1968	1968	1975	Données des auteurs pour 1976-1979
1931		0,9994	0,9987	0,9994	0,9994	0,9973
1947	10,35		0,9995	0,99999	0,99997	0,9989
1968	26,24	30,72		0,99996	0,9994	0,9986
1968	11,77	0,91	6,80		0,99999	0,9990
1975	12,38	1,97	88,82	7,35		0,9989
Données des auteurs pour 1976-1979	14,51	7,2	9,7	7,02	7,68	

* Au-dessus de la diagonale, on a porté les valeurs de l'indice de similitude ; sous la diagonale, les valeurs du critère d'identité [12], à $P = 0,05$, les valeurs critiques de $I = 7,81$ et à $P = 0,01$, $I = 11,34$.

génétique observée de la population ne peut être due à la diversité de la composition nationale des groupes étudiés. On peut supposer que l'hétérogénéité observée de la population de Moscou, surtout de la population russe, est en rapport avec l'arrivée de migrants d'origine russe venant de diverses régions de la R.S.F.S.R. différant du point de vue génétique.

La part des migrants venus à Moscou de la R.S.F.S.R. représente 72 % et celle des migrants venus d'Ukraine, 12 %. Le reste provient de toutes les autres régions de l'U.R.S.S. Dans les limites de la R.S.F.S.R., la plupart des migrants viennent de la région centrale (35,6 %) ; des autres régions, il en arrive de 1,5 à 6,5 % du nombre total [7]. Les données relatives à la composition génétique de la population russe de la partie européenne de la R.S.F.S.R. sont les plus exhaustives dans l'ouvrage des Bounak [4]. Cependant, à en juger d'après la description par l'auteur du schéma de dépouillement des premiers matériaux, l'idée de diversité génétique de la population que l'on peut tirer des données de l'ouvrage sera nettement minimisée.

Pour confronter le degré de similitude génétique entre la population russe de Moscou et la population des diverses régions de la partie européenne de la R.S.F.S.R., on a utilisé les données tirées de [4, 12] sur la base desquelles avaient été calculés les indices de similitude (*Tableau 4*). On peut suivre une augmentation nette de la simili-

Tableau 4

Indices de similitude entre les distributions des fréquences des groupes sanguins dans la population russe de Moscou et dans la population de certaines régions de la partie européenne de la R.S.F.S.R.
(d'après Oumnova et coauteurs [12] et Bounak [4])

Région	r	I	Région	r	I
d'Arkhangelsk de Vologda	0,9954	21,05	de Riazan' de Tambov	0,9995	9,36
de Perm' de Kirov	0,9946	136,25	de Lipetsk de Voronèje de Rostov		
de Pskov de Novgorod de Smolensk	0,9988	7,45	de Toula de Kalouga d'Orel	0,9999	0,93
de Kalinine d'Yaroslavl' de Kostroma d'Ivanovo de Vladimir	0,9997	8,34	de Briansk de Koursk Territoires de Krasnodar et de Stavropol'	0,9984	17,67

Remarque. A $P = 0,05$, $I = 7,81$,
à $P = 0,01$, $I = 11,34$.

tude génétique du nord vers Moscou de 0,9954 (régions d'Arkhangelsk et de Vologda) à 0,9997 (régions de la Haute Volga) et sa diminution vers le sud jusqu'à 0,9983 (Territoires de Krasnodar et de Stavropol'). On constate la même augmentation de la similitude génétique de l'ouest vers Moscou ($r = 0,9988$, régions de Novgorod, Smolensk et Pskov) et sa diminution en direction nord-est ($r = 0,9946$, régions de Perm' et de Kirov). Compte tenu des données mentionnées plus haut sur la structure démographique des couples conjugaux, il est évident que l'une des causes principales de l'hétérogénéité observée de la population de Moscou pour le locus ABO peut être la diversité des groupes comparés en ce qui concerne le lieu de naissance des sujets étudiés et de leurs parents. En effet, la comparaison des données démographiques obtenues par Kourbatova [8] et par nous-mêmes révèle une discordance évidente ($\chi^2 = 149,8$, $P < 0,001$). La structure matrimoniale des couples de parents des donneurs examinés

par nous en 1976 et en 1978-1979 (v. *Tableau 1*) diffère également ($\chi^2 = 18,5$, $P < 0,001$). En même temps, ces échantillons de donneurs se distinguent par les fréquences des groupes sanguins ($r = 0,9951$, $P < 0,05$). Cependant, si l'on ne compare les fréquences des groupes sanguins que parmi les sujets nés de mariages contractés entre Moscovites et de mariages contractés entre Moscovites et migrants venus de régions proches, on ne constate pas les différences signalées auparavant ($r = 0,9980$, $P = 0,50$). On tient compte également des sujets nés de mariages contractés entre Moscovites et migrants venus de régions limitrophes, ceci pour augmenter l'effectif de l'échantillon. Logiquement, cette réunion est tout à fait justifiée : la population de ces régions ne se distingue pas des Moscovites par les fréquences des distributions des groupes sanguins ou ne s'en distingue que fort peu (v. *Tableau 4*).

Les données examinées ci-dessus permettent de supposer une connexion entre la similitude génétique de la population de Moscou et de la population de différentes régions de la partie européenne de la R.S.F.S.R. et la part des migrants venus à Moscou de ces régions. Les coefficients de corrélation des rangs entre les indices de similitude des fréquences des groupes sanguins et la part des migrants venus à Moscou de différentes régions de R.S.F.S.R. confirment l'existence d'une telle connexion (*Tableau 5*).

Tableau 5

Coefficients de corrélation des rangs r_S entre les indices de similitude des distributions des fréquences des groupes sanguins ABO dans la population de Moscou et de certaines régions de la partie européenne de la R.S.F.S.R. et la part des migrants dans ces régions

Année de la recherche	r_S	Valeurs critiques de r_S
1958	0,86	$P_{0.05} = 0,79$
1970	0,93	$P_{0.01} = 0,94$
1978	0,79	

Remarque. La valeur de r_S calculée au total pour trois années d'observations au choix est égale à 0,85, la valeur critique de $r_{S\,0.01} = 0,56$.

Ainsi, les données démographiques et génétiques concernant la population de Moscou, ainsi que la population de certaines régions éloignées ou relativement proches de Moscou, permettent de considérer la migration comme le principal facteur déterminant l'hétérogénéité génétique de la population.

Les résultats obtenus montrent qu'il est nécessaire de recueillir des données démographiques au cours d'examens de la population urbaine au niveau de la génétique des populations ou médicale. En l'absence de données de ce genre, les résultats correspondants, même obtenus sur des échantillons à grand effectif, peuvent être aléatoires et ne refléter que des différences dans la composition démographique des groupes comparés.

RÉFÉRENCES

[1] *Алтухов Ю. П.* Генетический мониторинг популяций в связи с состоянием окружающей среды. Тезисы докладов XIV Международного генетического конгресса. Пленарные заседания. Симпозиумы. М., Наука, 1978, 16.

[2] *Altukhov Yu. P., Khilchevskaya R. I., Dubinin N. P.* Genetic monitoring of human populations in environment. In: Abstracts of V International Congress of Human Genetics. *Excerpta medica*, Mexico, 1976, 172.

[3] *Башлай А. Г.* Системы ABO, Rh и Kell-Cellano по данным о первичных донорах г. Москвы. Труды VII Международного конгресса антропологических и этнографических наук. 1, М., Наука, 1968, 491.

[4] *Бунак В. В.* Гено-географические зоны Восточной Европы, выделяемые по факторам крови ABO. *Вопр. антропол.*, 1969, 32, 6.

[5] *Дубинин Н. П., Алтухов Ю. П.* Мониторинг генетических последствий загрязнения окружающей среды. В кн.: Генетические последствия загрязнения окружающей среды. М., Мысль, 2, 14.

[6] *Dubinin N. P., Altukhov Yu. P.* Gene mutations (*de novo*) found in electrophoretic studies of blood proteins of infants with anomalous development. *Proc. Nat. Acad. Sci. U.S.A.*, 1979, 76, 5226.

[7] Итоги всесоюзной переписи населения 1970 года. 4, М., Статистика, 1974, 455 с.

[8] *Курбатова О. Л.* Опыт генодемографического исследования больших панмиксных популяций. Генетическая структура двух последовательных поколений жителей г. Москвы. *Вопр. антропол.*, 1975, 50, 30.

[9] *Прохоровская В. Д., Калабушкин Б. А.* Генетическая гетерогенность популяции г. Москвы по распределению групп крови. В кн.: Информационное обеспечение генетического мониторинга. Влияние загрязнения окружающей среды на здоровье населения в сельской местности. Ереван, Айастан, 1979, 72.

[10] *Рычков Ю. Г.* Сравнительное изучение генетического процесса в урбанизированной и изолированной популяциях. *Вопр. антропол.*, 1979, 63, 3.

[11] *Туманов А. К., Томилин В. В.* Наследственный полиморфизм изоантигенов и ферментов в норме и патологии человека. М., Медицина, 1969, 435.

[12] *Умнова М. А., Прокоп О., Пискунов Г. М., Сомусова Г. С., Учановская Т. А., Прозоровская Г. П.* Распределение различных факторов крови у населения Москвы. Труды VII Международного конгресса антропологических и этнографических наук. 1, М., Наука, 1968, 496.

[13] *Животовский Л. А.* Показатель сходства популяций по полиморфным признакам. *Ж. общ. Биол.*, 1979, 10, 4, 587.

TABLE DES MATIÈRES

Avant-propos 5

LA GÉNÉTIQUE DES POPULATIONS, L'ÉVOLUTION ET LE FOND GÉNIQUE DE LA BIOSPHÈRE

LA GÉNÉTIQUE BIOCHIMIQUE DES POPULATIONS ET L'ÉVOLUTION
par *Y. Altoukhov* 9

L'ORGANISATION SYSTÉMIQUE DES POPULATIONS ET SON IMPORTANCE POUR LA SAUVEGARDE ET L'UTILISATION RATIONNELLE DU FOND GÉNIQUE DE LA BIOSPHÈRE
par *Y. Altoukhov* et *E. Salmenkova* 35

VARIABILITÉ DANS UNE POPULATION ACTUELLE ET DANS UNE POPULATION DE L'HOLOCÈNE MOYEN DE *LITTORINA SQUALIDA*
par *B. Kalabouchkine* 61

SIMULATION EXPÉRIMENTALE DES PROCESSUS GÉNÉTIQUES DANS UN SYSTÈME DE POPULATIONS DE *DROSOPHILA MELANOGASTER* CORRESPONDANT AU MODÈLE ÉCHELONNÉ CIRCULAIRE. I
par *Y. Altoukhov, A. Bernachevskaïa, A. Milichnikov* et *T. Novikova* 72

SIMULATION EXPÉRIMENTALE DES PROCESSUS GÉNÉTIQUES DANS UN SYSTÈME DE POPULATIONS DE *DROSOPHILA MELANOGASTER* CORRESPONDANT AU MODÈLE ÉCHELONNÉ CIRCULAIRE. II
par *Y. Altoukhov* et *A. Bernachevskaïa* 84

MÉTHODE D'IDENTIFICATION DES PHÉNOTYPES MORPHOLOGIQUEMENT « MOYENS » ET « EXTRÊMES » D'APRÈS L'ENSEMBLE DES CARACTÈRES QUANTITATIFS
par *L. Jivotovski* et *Y. Altoukhov* 94

EFFETS DE LA SÉLECTION MODALE ET DIRIGÉE POUR L'ENSEMBLE DES CARACTÈRES SUR LE COTONNIER *GOSSYPIUM HIRSUTUM* L.
par *Y. Altoukhov, L. Jivotovski, S. Sadykov, B. Abdoullaev* et *B. Kalabouchkine* 100

PARTICULARITÉS DU DESSIN DU CARACUL ET STRUCTURE GÉNÉTIQUE DES GROUPES DE MOUTONS DE KARA-KOUL' RAPPORTÉS AUX TYPES MORPHOLOGIQUEMENT « MOYENS » ET « EXTRÊMES »
par *Y. Altoukhov, N. Sarsenbaev, K. Afanassiev, T. Malinina, A. Milichnikov* et *E. Pobédonostséva* 106

MESURES DE LA VARIABILITÉ DES POPULATIONS SELON LES CARACTÈRES POLYMORPHES
par *L. Jivotovski* 122

MESURES DE LA VARIABILITÉ DES POPULATIONS SELON UN ENSEMBLE DE CARACTÈRES QUANTITATIFS
par *L. Jivotovski* 130

L'ENVIRONNEMENT ET LE FARDEAU GÉNÉTIQUE DES POPULATIONS HUMAINES

PROBLÈMES GÉNÉTIQUES DE LA POLLUTION DE L'ENVIRONNEMENT
par *R. Khiltchevskaïa* 151

MONITORING GÉNÉTIQUE DES POPULATIONS HUMAINES EN ACCORD AVEC L'ÉTAT DE L'ENVIRONNEMENT
par *Y. Altoukhov* 168

SÉLECTION DIRIGÉE CONTRE LES VARIANTES ÉLECTROPHORÉTIQUES RARES DES PROTÉINES ET CADENCES DE LA MUTATION SPONTANÉE DANS LES POPULATIONS
par *Y. Altoukhov, V. Doukharev* et *L. Jivotovski* 191

APPROCHE DU PROBLÈME DE LA RÉSISTANCE BIOLOGIQUE NON SPÉCIFIQUE DE L'ORGANISME HUMAIN AU NIVEAU DE LA GÉNÉTIQUE DES POPULATIONS. I
par *Y. Altoukhov, O. Botviniev* et *O. Kourbatova* 208

APPROCHE DU PROBLÈME DE LA RÉSISTANCE BIOLOGIQUE NON SPÉCIFIQUE DE L'ORGANISME HUMAIN AU NIVEAU DE LA GÉNÉTIQUE DES POPULATIONS. II
par *O. Botviniev, O. Kourbatova* et *Y. Altoukhov* 218

RÉSISTANCE NON SPÉCIFIQUE DE L'ORGANISME ET PRÉDISPOSITION À LA LEUCÉMIE AIGUË PARMI LES ENFANTS
par *O. Kourbatova, O. Botviniev, M. Dechtchokina, I. Torbiak* et *Y. Altoukhov* 232

NIVEAUX DE POLYMORPHISME ET D'HÉTÉROZYGOTIE DE LA POPULATION RUSSE DE MOSCOU : DONNÉES RELATIVES À 22 LOCI GÉNIQUES CODANT POUR LES PROTÉINES SANGUINES
par *Y. Altoukhov, R. Khiltchevskaïa* et *A. Chourkhal* 244

LES MARQUEURS GÉNIQUES ET LES MALADIES : PARTICULARITÉS GÉNÉTIQUES, ANTHROPOMÉTRIQUES ET CLINIQUES DES ENFANTS ATTEINTS DE PNEUMONIE AIGUË
par *Y. Altoukhov, O. Kourbatova, O. Botviniev, K. Afanassiev, T. Malinina, O. Kholod, L. Strelkova* et *N. Ivanova* 256

LA MIGRATION COMME FACTEUR D'HÉTÉROGÉNÉITÉ GÉNÉTIQUE DE LA POPULATION DES GRANDES VILLES
par *B. Kalabouchkine, V. Prokhorovskaïa, R. Khiltchevskaïa* et *Y. Altoukhov* 271

Imprimé en Union Soviétique